

Entwicklung eines adaptiven Hilfesystems für multimodale Anzeige-Bedienkonzepte im Fahrzeug



Inaugural-Dissertation zur Erlangung der Doktorwürde der
Philosophischen Fakultät IV (Sprach- und
Literaturwissenschaften) der Universität Regensburg

vorgelegt von
Alexander Hof

aus
Regensburg
2007

Erstgutachter: Prof. Dr. phil. Christian Wolff

Zweitgutachter: Prof. Dr. rer. soc. Rainer Hammwöhner

Danksagung

Die vorliegende Arbeit entstand an der Professur für Medieninformatik (Prof. Dr. Christian Wolff) der Universität Regensburg in Zusammenarbeit mit dem Fachbereich Entwicklung Innenraum/Fahrerarbeitsplatz der BMW Group München.

Mein besonderer Dank gilt in erster Linie meinen Betreuern: Professor Wolff für seine immerwährende Gesprächsbereitschaft, die stets konstruktiven Beiträge und den nötigen Freiraum, der für eine Industriepromotion nötig ist; Dr. Alexander Huber (BMW Group) für seine stets wertvollen fachlichen Beiträge, viele Diskussionen und die tatkräftige Unterstützung in organisatorischen Belangen; Dr. Eli Hagen (BMW Group) für die Betreuung in fachlicher und wissenschaftlicher Hinsicht sowie die Bereitschaft, Konzepte kritisch zu hinterfragen.

Dank gebührt auch Constanze Schwarz und Dr. Robert Neuss (Usaneers GmbH), ohne deren Hilfe die Implementierung des Prototyps nicht möglich gewesen wäre. Ferner danke ich Professor Rainer Hammwöhner, Dr. Ludwig Hitzenberger und den Mitgliedern des Doktorandenseminars für fruchtbare Diskussionen und hilfreiche Kommentare. Worte des Dankes gehen auch an Stefan Sagstetter, der mich im Rahmen seiner Masterarbeit bei der Planung und Durchführung der Evaluierung unterstützte, sowie an Heike Kramer, die meine Arbeit korrektur gelesen hat.

Für den Rückhalt – v.a. in motivationsdefizitären Zeiten – danke ich meiner Freundin Nadine, meinen Eltern Evelyn und Hans, meinem Bruder Michael und meinen Freunden.

Beatum non eum esse quem vulgus appellat, ad quem pecunia magna confluxit, sed illum cui bonum omne in animo est, erectum et excelsum et mirabilia calcantem, qui neminem videt cum quo se commutatum velit, qui hominem ea sola parte aestimat qua homo est.

Lucius Annaeus Seneca, Epistulae morales ad Lucilium

Inhaltsverzeichnis

1	Einleitung	13
2	Grundlagen sprachverarbeitender Dialogsysteme	17
2.1	Mensch-Maschine-Interaktion	17
2.1.1	Mensch-Maschine-System und Mensch-Maschine-Schnittstelle	18
2.1.2	Multimodale Mensch-Maschine-Schnittstellen	20
2.1.3	Adaptive Mensch-Maschine-Schnittstellen und Benutzermodellierung	23
2.2	Sprachdialogsysteme	24
2.2.1	Evolution von Sprachdialogsystemen	24
2.2.2	Architektur eines Sprachdialogsystems	25
2.2.3	Vorteile sprachverarbeitender Systeme	28
2.3	Gestaltungsaspekte beim Entwurf von Sprachdialogsystemen . . .	28
2.3.1	Softwareergonomische Richtlinien	28
2.3.2	Kognitionspsychologische Aspekte	29
2.4	Fazit	32
3	Stand der Technik im Bereich interaktiver Hilfesysteme	35
3.1	Definition und Klassifikation von Hilfesystemen	35
3.2	Konzepte interaktiver Hilfesysteme	37
3.2.1	COMFOHELP	37
3.2.2	PHI	39
3.2.3	Lumière (Microsoft Office Assistant)	40
3.2.4	Kyoto Sightseeing Guide	41
3.2.5	SmartAidè	43
3.2.6	Targeted Help	45
3.2.7	MATCH: Multimodal Access to City Help	47
3.2.8	Hilfesysteme für SDS in automotiven Umgebungen	49
3.3	Fazit	53
4	Analyse und Bewertung interaktiver Hilfesysteme	57
4.1	Generelle Anforderungen an Hilfesysteme künftiger Fahrzeuggenerationen	57
4.2	Bewertung bestehender Hilfesysteme in Fahrzeugen	60
4.3	Bewertung alternativer Hilfekonzepte	65

4.4	Fazit	67
5	Modellierung eines adaptiven, multimodalen Hilfesystems	69
5.1	Allgemeine Hilfe	69
5.2	Kontextspezifische Hilfe	70
5.2.1	Antizipative Reorganisation der Hilfeinhalte durch Analyse des Benutzerverhaltens	72
5.2.2	Generische Anpassung der Dialogstrategie an Nutzererfahrung	79
5.2.3	Visuelle Repräsentation der allgemeinen und kontext- spezifischen Hilfe	86
5.3	Kontextunabhängige, natürlichsprachliche Hilfe	90
5.4	Experimentelle Bestimmung der Systemparameter	95
5.4.1	Lernen vom Sprachkommandos	95
5.4.2	Ausbildung von Transferwissen	97
5.4.3	Vergessen von Sprachkommandos	98
5.5	Auswahl der Beschreibungsmodelle	100
5.6	Fazit	102
6	Implementierung des Prototyps	105
6.1	Systemarchitektur der Dialogsimulation	105
6.2	Kontextabhängige Hilfe	107
6.3	Kontextunabhängige Hilfe	110
6.4	Fazit	112
7	Evaluierung des Hilfesystems	113
7.1	Arbeitshypothesen	113
7.2	Testdesign	114
7.3	Testplan	114
7.3.1	Versuchspersonen	114
7.3.2	Versuchsaufbau	115
7.3.3	Versuchsablauf	117
7.3.4	Aufgabenstellung	118
7.4	Ergebnisse	119
7.4.1	Bewertung allgemeiner Systeminteraktionen	119
7.4.2	Analyse des Blickverhaltens	122
7.4.3	Subjektive Bewertung der Hilfesysteme	123
7.4.4	Ermittlung der Usability anhand des Evaluierungsframe- works PARADISE	126
7.4.5	Messung mentaler Beanspruchung mittels NASA TLX . .	135
7.5	Diskussion der Ergebnisse	137
7.6	Fazit	140
8	Zusammenfassung und Ausblick	143

8.1	Anpassung der Hilfeinhalte und der Dialogstrategie an das Wissen und den Erfahrungsgrad des Benutzers	144
8.2	Multimodale Hilfefunktionalität	145
8.3	Kontextunabhängige Hilfefunktionalität	145
8.4	Evaluierung des Hilfesystems	145
8.5	Ausblick	146
Literaturverzeichnis		149
A Abkürzungsverzeichnis		161
B Intentionen Pool		165
C Testdrehbuch		167
D Fragebogen		175

Abbildungsverzeichnis

1.1	Cockpits im BMW 328 und im Mercedes SL	14
1.2	Cockpits im Audi Q7 und BMW 6er	14
2.1	IFIP-Modell für Mensch-Maschine-Systeme	18
2.2	Seeheim Architekturmodell für Mensch-Maschine-Systeme	19
2.3	Ueware Begriff in Anlehnung an Reuther (2003)	20
2.4	Multimodale Schnittstellen	21
2.5	Vereinfachtes Klassifikationsschema multimodaler Systeme nach Nigay und Coutaz (1993)	21
2.6	Architekturmodell für Sprachdialogsysteme nach Keller (2004) und McTear (2002)	26
2.7	Grundprinzip von Hidden-Markov Modellen	26
2.8	Ergonomische Anforderungen für Bürotätigkeiten mit Bildschirmgeräten Teil 10: Grundsätze der Dialoggestaltung (EN ISO 9241-10)	29
2.9	Acht Goldene Regeln des Interface Designs nach Shneiderman (1997)	30
2.10	Richtlinien für die Gestaltung von Sprachdialogen	30
2.11	Serielle Positionskurve	32
3.1	Domain Concept Tree	42
3.2	Schematische Darstellung des DCT	43
3.3	Aufbau des Targeted Help Moduls	45
3.4	MATCH System	48
3.5	GUIs und Bedienelemente verschiedener Anzeige-Bedienkonzepte .	50
3.6	Klassifikation von Hilfesystemen	53
3.7	Klassifikation interaktiver Hilfesysteme	54
3.8	Klassifikation bestehender Hilfesysteme für SDS in Fahrzeugen . .	55
4.1	Kenntnisse eines wenig erfahrenen Benutzers über Systemfunktionen	58
4.2	Kenntnisse eines erfahrenen Benutzers über Systemfunktionen . .	59
4.3	Anforderung an multimodale Dialogsysteme im Kraftfahrzeug . .	59
4.4	Bewertung von Hilfesystemen in Fahrzeugen	64
5.1	Ausnutzung des Primacy- und Recency-Effekts	71
5.2	Systemmodell des Benutzers beim Erstkontakt	72
5.3	Abstrakte Struktur des Hilfesystems beim Erstkontakt mit dem SDS	73
5.4	Veränderte Anordnung der Sprachkommandos	74

5.5	Exemplarischer Verlauf einer Lernkurve	75
5.6	Parameter der Funktionen für die Hilfeausgabe	76
5.7	Exemplarischer Verlauf einer Vergessenskurve	78
5.8	Exemplarischer Verlauf von i_A	80
5.9	Schematische Darstellung der Hilfetafel	87
5.10	Visuelle Unterstützung der Hilfe	88
5.11	Überlappung multimodaler Ein- und Ausgaben	89
5.12	Disambiguierungstafel	90
5.13	Zyklischer, ungerichteter Graph	92
5.14	Algorithmus von Dijkstra zur Bestimmung des kürzesten Pfades zwischen zwei Knoten.	93
5.15	Anweisung zum Thema „Eintrag aus dem Telefonbuch löschen“ . .	94
5.16	Resultierende Lernkurve und Lernkurve basierend auf einer Potenz- funktion (gestrichelte Linie).	97
5.17	Piktogramme zur Bildung der Intentionen 'Telefon', 'Autobahn ver- meiden' und 'Radio'	99
5.18	Resultierende Vergessenskurve und Vergessenskurve basierend auf einer Exponentialfunktion	100
5.19	Berechnung des Schwellenwerts N	101
5.20	Verlauf der Indices einer Funktion f_1 sowie der Funktion f_ϕ einer Funktionsfamilie ϕ	102
6.1	Systemarchitektur des Sprachdialogsystems	106
6.2	Systemarchitektur des visuell-haptischen Bediensystems	106
6.3	Ermittlung der Zeit bei der ersten Verwendung einer Funktion . .	108
6.4	Klassifikation von Prototypen nach Nielsen (1994)	111
7.1	Übersicht über die Probanden	115
7.2	Versuchsaufbau Fahrsimulator	116
7.3	Versuchsaufbau im Fahrzeug	117
7.4	Gruppierung der Probanden	119
7.5	Erkennerfehler (Mittelwert)	120
7.6	Aufgabendauer (Median)	121
7.7	Sprecherbeiträge (Mittelwert)	121
7.8	Optionsaufrufe (Mittelwert)	122
7.9	Analyse des Blickverhaltens	123
7.10	Veränderung der Dialoge	124
7.11	Länge der Hilfeausgaben	124
7.12	Unklarheiten nach Verwendung der Hilfe	125
7.13	Visuelle Hilfe	125
7.14	Einsatz des Controllers	126
7.15	Wie nützlich ist für Sie diese Art der [kontextunabhängigen] Hilfe?	126
7.16	Grundlagen des PARADISE Frameworks	127

7.17	Konfusionsmatrix	128
7.18	Konfusionsmatrtix nach Hassel (2006)	130
7.19	Nutzerzufriedenheitsmaße	131
7.20	Aufgabenerfolg κ^*	133
7.21	Erfolgsmaß (Nutzer)	133
7.22	Erfolgsmaß (System)	134
7.23	Bewertungsskala des NASA-TLX	136
7.24	Mentale Belastung basierend auf dem NASA-TLX	137

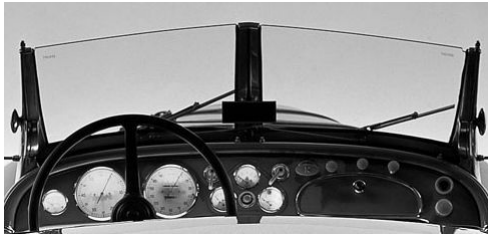
Kapitel 1

Einleitung

Elektronische Fahrerassistenz- und Fahrerinformationssysteme (FAS bzw. FIS) gehören in vielen Fahrzeugen mittlerweile zur Grundausstattung. Es werden jedoch stets neue Komponenten in die Fahrzeuge integriert, um dem Fahrer höchste Sicherheit und maximalen Komfort zu garantieren. Automatische Abstandskontrollsysteme sollen den Fahrer bei Folgefahrten im fließenden Verkehr unterstützen, im Kommunikationsbereich wird der Einbindung alltäglich gewordener Kommunikationsmittel wie Mobiltelefonie, Kurznachrichten (SMS) oder E-Mail ebenso große Bedeutung beigemessen wie der Integration diverser Online-Dienste. Auf dem Gebiet der Entertainmentsysteme wird die Anbindung von digitalen Komponenten und Technologien wie z.B. digitales Fernsehen und Radio vorangetrieben.

Das steigende Angebot an Funktionen bringt einerseits viele Annehmlichkeiten für den Fahrer mit sich, jedoch ergeben sich daraus auch neue Ansprüche bezüglich der Bedienbarkeit dieser Funktionen. Die Primäraufgabe des Fahrens darf durch Sekundäraufgaben, die durch die Bedienung verschiedener FIS, FAS und Komfortsysteme entstehen, nicht in Mitleidenschaft gezogen werden (Totzke, 2001; Piechulla et. al., 2002). Aus diesem Grund muss die Interaktion mit derartigen Systemen eine anwenderfreundliche Bedienung und minimales Ablenkunspotenzial in sich vereinen. Eine weitere zentrale Herausforderung besteht darin, alle Funktionen und Geräte im zur Verfügung stehenden Bauraum unterzubringen. Die bloße Hinzunahme von Bedienelementen würde schnell zu einer Überfrachtung des Fahrerarbeitsplatzes führen.

Aus diesem Grund ist es nötig, neue Bedienparadigmen zu integrieren, um dem Fahrer möglichst viele Funktionalitäten anzubieten, und ihm gleichzeitig eine Methode an die Hand zu geben, die Funktionalitäten einfach, effizient und sicher zu verwalten. Vor diesem Hintergrund entwickelte man bei diversen Automobilherstellern neue Bedienkonzepte, bei deren Entwurf man Schritt für Schritt von der ursprünglichen Formel „Gerät = Funktion + Bedienung“ (vgl. Abbildung 1.1(a)) abwich. Die Integration von graphischen Bedienoberflächen ermöglichte erstmals eine teilweise Mehrfachbelegung einzelner Eingabeelemente (vgl. Abbildung 1.1(b)).



(a) Cockpit im BMW 328 (Bj. 1936).
Quelle: BMW Group (2007)



(b) Mittelkonsole im Mercedes SL (Bj. 2005).
Quelle: Mercedes Benz (2007)

Abbildung 1.1: Für jede Funktion im BMW 328 existiert ein eigener Dreh- oder Drücksteller sowie ein separates Anzeigeelement; Die Integration eines Displays im Mercedes SL optimiert die Bedienbarkeit.

Den entscheidenden Schritt bei der Entwicklung von richtungsweisenden Anzeige-Bedien-Konzepten unternahmen die Hersteller Audi und BMW, welche im Jahr 2001 die Bedienkonzepte *Multimedia Interface* (Audi) bzw. *iDrive* (BMW) in den jeweiligen Oberklasse-Limousinen vorstellten. Beiden Konzepten ist gemein, dass die Fahrfunktionen von den Komfortfunktionen getrennt sind. Während die Fahrfunktionen wie z.B. Hebel für Scheibenwischer, Blinker und beim BMW auch die Schaltung rund um das Lenkrad positioniert sind, befinden sich die Bedienelemente der Komfortfunktionen wie z.B. die Regler der Klimaautomatik sowie Fahrerassistenz- und Fahrerinformationssysteme zentral im optimalen Greifraum des Fahrers und können zudem vom Beifahrer bedient werden. Neben der ergonomisch vorteilhafteren Eingabe ermöglicht das zugehörige Display



(a) Cockpit im Audi Q7 (Bj. 2006).
Quelle: Audi (2007)



(b) Cockpit im BMW 6er (Bj. 2005).
Quelle: BMW Group (2007)

Abbildung 1.2: Cockpits im Audi Q7 und BMW 6er

auf auf der Mittelkonsole eine bessere Ablesbarkeit von Informationen (Bengler et. al., 2002).

Obgleich die Bedienkonzepte so ausgelegt sind, dass sie intuitiv anwendbar sind, besitzt die visuell-haptische Bedienung über zentrale Bedienelemente und Displays den Nachteil, dass der Fahrer während der Eingaben eine Hand vom Lenkrad nehmen und den Blick von der Straße abwenden muss. Um dem Fahrer größtmögliche Sicherheit bei maximalem Komfort zu bieten, stellen Audi, BMW und Mercedes zusätzlich zur visuell-haptischen Bedienung ein Sprachdialogsystem (SDS) zur Verfügung.

Die Integration eines SDS im Fahrzeug birgt zwei entscheidende Pluspunkte in sich: Zum einen ist Sprache eine natürliche und intuitive Kommunikationsform. Der Umgang mit Sprache muss nicht explizit erlernt werden, was bedeutet, dass die Sprachbedienung im Vergleich zu anderen Bedienkonzepten leichter zu erlernen ist (Wahlster, 1999). Zum anderen belegen Studien, dass die Verarbeitung akustischer Informationen im Fahrzeug Ablenkungspotenziale reduziert (Bengler, 1995). Im Vergleich zu visuellen oder haptischen Bediensystemen wirken sich nach Vollrath und Totzke (2000) akustische Bedienkonzepte nicht negativ auf die Fahrzeugführung (Längs- und Querführung) aus und reduzieren zudem die körperliche und mentale Belastung (Erke et. al., 2001). Aus diesem Grund scheint die Implementierung eines SDS im Fahrzeug nicht nur sinnvoll, sondern notwendig.

Da die Bedienung per Sprache jedoch relativ neu und für viele Fahrer auch ungewohnt ist, ist es wichtig, den Einstieg in die Sprachbedienung möglichst einfach zu gestalten. Einen entscheidenden Beitrag für die intuitive Bedienung leistet das Konzept „speak what you see“, d.h. der Benutzer kann jederzeit die Kommandos sprechen, die auf dem Display zu sehen sind (Hagen et. al., 2004). Für den Fall, dass eine gewünschte Funktion nicht angezeigt wird oder der zugehörige Befehl nicht bekannt ist, muss ein leistungsstarkes Hilfesystem zur Verfügung stehen, das den Benutzer bei der Erreichung seines Dialogziels unterstützt.

Ziel der vorliegenden Arbeit ist es, ein Hilfefunktion zu entwickeln, das vorhandene Schwachstellen existierender Hilfesysteme beseitigt und die steigenden Funktionsumfänge in künftigen Fahrzeugbaureihen beherrschbar macht. Von zentraler Bedeutung ist dabei die Anpassung der Hilfeausgaben an das Wissen und die Erfahrungen des Benutzers sowie die Ergänzung der Hilfe um visuelle und haptische Ein- und Ausgabemöglichkeiten.

Eine detaillierte Beschreibung der Konzepte hinsichtlich der kontextspezifischen und kontextunabhängigen Helfefunktionalität erfolgt in Kapitel 5. Die theoretischen Grundlagen aus dem Bereich der Mensch-Maschine-Interaktion sowie essenzielle Gestaltungsaspekte für den Entwurf von Sprachdialogsystemen werden in Kapitel 2 erläutert. Der Weiterentwicklung bestehender Hilfesysteme geht zudem eine Standortbestimmung im Bereich interaktiver Hilfesysteme voraus. In Kapitel 3 wird der Status Quo interaktiver Hilfesysteme aus unterschiedlichen Anwen-

dungsdomänen präsentiert, Kapitel 4 beinhaltet eine Analyse dieser Hilfekonzepte. Um die in Kapitel 5 aufgeführten Maßnahmen zur Verbesserung des Hilfesystems erlebbar zu machen, wird ein auf dem iDrive-Nachfolgesystem basierender Prototyp implementiert, der alle konzeptuellen Ansätze in sich vereinigt (Kapitel 6). Dieser Prototyp wird dem Hilfesystem des aktuellen SDS in einem Vergleichstest gegenübergestellt, dessen Testdesign neben den Ergebnissen der Evaluierung in Kapitel 7 erläutert wird. Eine Zusammenfassung der Kernaspekte dieser Arbeit sowie ein Ausblick auf künftige Forschungsthemen im Bereich adaptiver Hilfesysteme erfolgt in Kapitel 8.

Kapitel 2

Grundlagen sprachverarbeitender Dialogsysteme

In diesem Kapitel werden zunächst grundlegende Begriffe definiert, die für das Verständnis dieser Arbeit eine zentrale Rolle spielen. Nach einer kurzen Erläuterung elementarer Begriffe aus dem Bereich der Mensch-Maschine-Interaktion (MMI) erfolgt eine eingehendere Betrachtung multimodaler und adaptiver MMI. Neben einem allgemeinen Überblick über softwareergonomische und psychologische Themengebiete, die einen wichtigen Beitrag zur Entwicklung von Mensch-Maschine-Systemen leisten, werden kurz die Grundlagen der maschinellen Sprachverarbeitungssysteme skizziert. Abschließend erfolgt eine Klassifikation von Hilfesystemen sowie die Bestimmung des Status Quo im Bereich sprachbasierter Hilfesysteme.

2.1 Mensch-Maschine-Interaktion

Der Begriff der Mensch-Maschine-Interaktion (MMI) wird in verschiedenen Kontexten unterschiedlich beschrieben. MMI ist eine interdisziplinär ausgerichtete Forschungsrichtung und bewegt sich im Spannungsfeld von Informatik, Psychologie, Design und Ergonomie (Helander et. al., 1997). Mit besonderem Fokus auf die Sprachbedienung findet auch die Linguistik Eingang in die MMI. Eine abstrakte Definition des Begriffs der MMI liefert Hewitt et. al. (1992):

Human-Computer-Interaction is a discipline concerned with the design, evaluation and implementation of interactive computing systems for human use and with the study of major phenomena surrounding them.

Die besondere Herausforderung auf dem Gebiet der MMI besteht darin, die einzelnen Teildisziplinen so miteinander zu verknüpfen, dass die Interaktion zwischen Mensch und Maschine möglichst optimal auf die Bedürfnisse und Fähigkeiten des Menschen abgestimmt ist.

2.1.1 Mensch-Maschine-System und Mensch-Maschine-Schnittstelle

Innerhalb der MMI bezeichnet der Begriff Maschine technische Systeme unterschiedlichster Art, z.B. Software-Systeme, Fahrzeuge etc. Daran angelehnt kann ein Mensch-Maschine-System als Regelkreis verstanden werden, bei dem Mensch und Maschine ein geschlossenes System bilden, um ein bestimmtes Ziel zu erreichen. Eine allgemeine Definition des Begriffs Mensch-Maschine-System (MMS) wird in Geiser (1990) und Johannson (1993) gegeben:

Bei einem Mensch-Maschine-System wirkt der Mensch mit einer Maschine mit dem Ziel zusammen, eine selbstgewählte oder vorgegebene Aufgabe zu lösen.

Im Jahr 1983 entwarf Dzida das sog. IFIP-Modell (International Federation for Information Processing), mit dessen Hilfe die allgemeine Struktur eines Mensch-Maschine-Systems beschrieben werden konnte. Das IFIP-Modell definiert vier Schnittstellen (Ein- & Ausgabeschnittstelle, Dialogschnittstelle, Werkzeugschnittstelle, Organisationsschnittstelle; siehe Abbildung 2.1). Neben dem organisatorisch

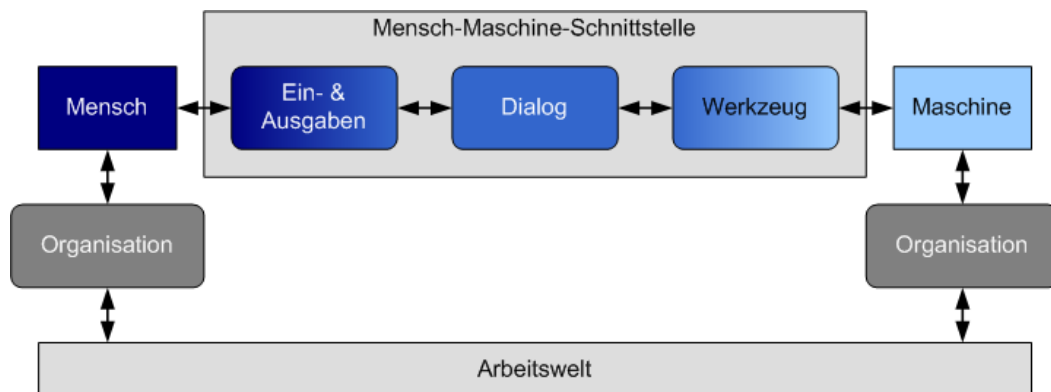


Abbildung 2.1: IFIP-Modell für Mensch-Maschine-Systeme

orientierten IFIP-Modell wurde das sog. Seeheim-Modell entwickelt (siehe Abbildung 2.2) (Green, 1985). Es beschreibt ein Architekturmodell zur einheitlichen Realisierung von Mensch-Maschine-Systemen, wobei eine Dreiteilung der Mensch-Maschine-Schnittstelle in drei Schichten (Präsentation, Dialogmanagement, Applikationsinterface; siehe Abbildung 2.2) postuliert wird.

Im Kern weisen beide Modelle ähnliche Komponenten auf, wobei der Austausch von Informationen zwischen Mensch und Maschine stets über eine Mensch-Maschine-Schnittstelle erfolgt. Nach Baggen und Hemmerling (2000) versteht man unter einer Mensch-Maschine-Schnittstelle die

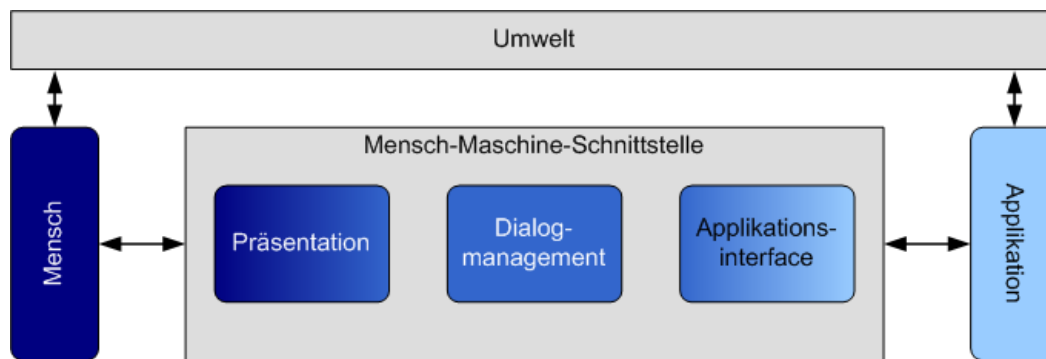


Abbildung 2.2: Seeheim Architekturmodell für Mensch-Maschine-Systeme

[...] Komponenten eines Mensch-Maschine-Systems, die für die Interaktion mit dem Benutzer relevant sind.

Abgeleitet aus den vorhandenen Modellen besteht eine Mensch-Maschine-Schnittstelle aus folgenden Komponenten:

- Präsentationskomponente: umfasst alle statischen Aspekte einer Bedienoberfläche und bietet dem Benutzer verschiedene Ein- und Ausgabemöglichkeiten.
- Dialogkomponente: beinhaltet die dynamischen Elemente zur Benutzerführung. Die Ein- und Ausgaben des Benutzers werden an dieser Stelle verwaltet.
- Interfacekomponente: Interaktionen zwischen Mensch und Maschine werden in maschinenverständlichen Beschreibungsformaten bearbeitet, z.B. VoiceXML.

Obgleich die bisher genannten Modelle eine allgemeine Beschreibung von Mensch-Maschine-Systemen implizieren, sind diese noch sehr stark in der Interaktion zwischen Mensch und Computer verankert. Gerade die Entwicklung von Bediensystemen z.B. in der Produktionstechnik oder auch im Automobilbereich lässt jedoch die Grenzen zwischen Hard- und Softwarebediensystemen verschwimmen. Aus diesem Grund wurde für alle der Interaktion dienenden Komponenten eines Mensch-Maschine-Systems der Begriff Useware eingeführt, der sich nach Zühlke (2002) wie folgt definiert:

Useware umfasst alle der Nutzung einer Maschine oder Anlage dienenden Hard- und Software-Komponenten.

Useware bildet die Schnittstelle zwischen Hardware, Software und dem Benutzer (siehe Abbildung 2.3), wobei Useware sich stets an den Fähigkeiten und Limitationen des Benutzers orientiert. Verallgemeinert man die Anwendungsdomänen für das IFIP- und Seeheim-Modell auf alle technischen Systeme und erweitert

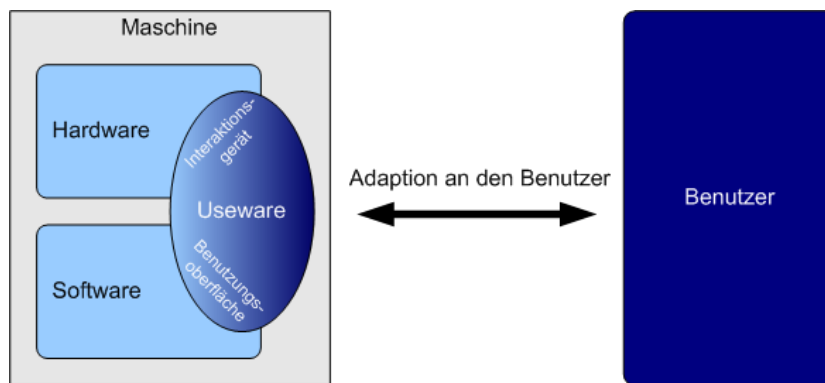


Abbildung 2.3: Useware Begriff in Anlehnung an Reuther (2003)

dementsprechend die möglichen Ein- und Ausgabekomponenten, so behalten diese Modelle auch im Hinblick auf Useware ihre Gültigkeit.

2.1.2 Multimodale Mensch-Maschine-Schnittstellen

Wie in Kapitel 2.1.1 beschrieben, muss Useware sich an die Fähigkeiten eines Benutzers anpassen. Während beispielsweise im Fahrzeug vornehmlich Eingaben haptisch mittels Dreh- oder Drückreglern und Ausgaben über visuelle Feedbacksysteme erfolgen, sollen optimierte Bediensysteme multimodale Interaktionen erlauben. Der Begriff Multimodalität beschreibt im Kontext der Mensch-Maschine-Interaktion die Interaktion eines Benutzers mit der Maschine über verschiedene Sinnesmodalitäten und definiert sich nach Benoit (1997) wie folgt:

Multimodal systems represent and manipulate information from different human communication channels at multiple levels of abstraction. Multimodal systems can automatically extract meaning from multimodal, raw input data, and, conversely they produce perceivable information from symbolic abstract representations.

Je nachdem ob eine Modalität zur Ein- oder Ausgabe verwendet wird, spricht man von Aktionsmodalität oder Wahrnehmungsmodalität. Unter Berücksichtigung der menschlichen Wahrnehmungs- und Aktionsmodalitäten, ergeben sich nach Hedicke (2002) drei mögliche Schnittstellen: eine auditive, eine visuelle und eine haptische (siehe Abbildung 2.4). Über jede dieser Schnittstellen können Ein- und Ausgaben erfolgen.

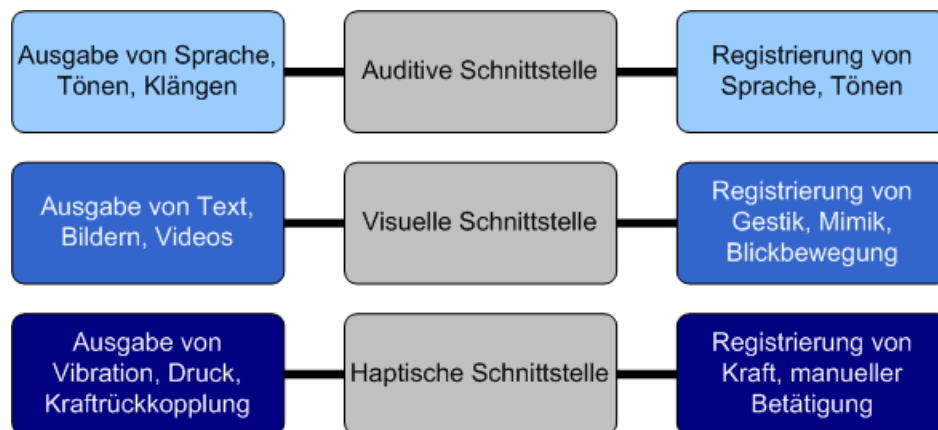


Abbildung 2.4: Multimodale Schnittstellen

2.1.2.1 Klassifikation multimodaler Systeme

Nigay und Coutaz (1993) klassifizieren multimodale Systeme anhand des zeitlichen Gebrauchs von multimodalen Ein- und Ausgaben, der Fusion von Daten verschiedener Ein- und Ausgabegeräte und des Abstraktionsgrads der Informationsinterpretation. Da die Informationsinterpretation im Rahmen dieser Arbeit stets auf semantischem Niveau erfolgt, kann die zeitliche Kennzeichnung vernachlässigt und das Klassifikationsschema vereinfacht dargestellt werden (siehe Abbildung 2.5). Der Gebrauch von Modalitäten ist aufgeteilt in sequentielle und paral-

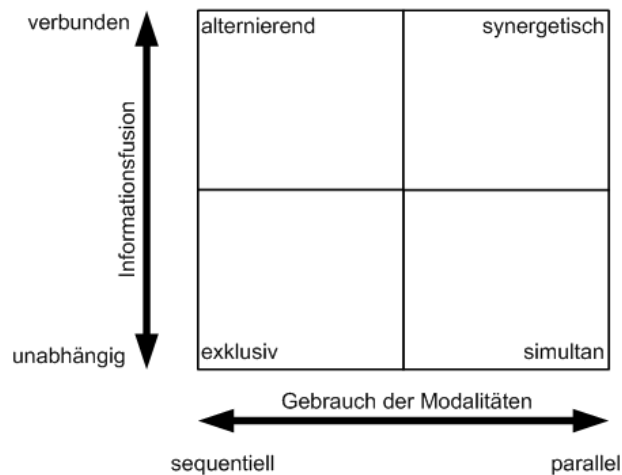


Abbildung 2.5: Vereinfachtes Klassifikationsschema multimodaler Systeme nach Nigay und Coutaz (1993)

lele Interaktionen. Paralleler Gebrauch erlaubt dem Benutzer, mehrere Modalitäten gleichzeitig zu verwenden, wohingegen beim sequentiellen Gebrauch Interak-

tionen lediglich Schritt für Schritt durchgeführt werden können. Der Grad der Informationsfusion gibt an, inwiefern Informationen aus verschiedenen Modalitäten miteinander verbunden werden oder nicht.

Brey und Salmen (2003) verfeinern dieses Klassifikationsschema im Hinblick auf die Informationsfusion, indem zusätzlich unterschieden wird, ob Ein- oder Ausgaben komplementär zueinander erfolgen oder redundant sind. Somit ergibt sich folgendes Klassifikationsschema für alternierende bzw. synergetische Systeme:

- (Unimodale Interaktion: Interaktion in einer Modalität.)
- Sequentiell-komplementäre Interaktion: verschiedene Informationen werden in verschiedenen Modalitäten zeitlich aufeinanderfolgend verarbeitet.
- Sequentiell-redundante Interaktion: identische Informationen werden zeitlich versetzt und über verschiedene Modalitäten verarbeitet.
- Simultan-komplementäre Interaktion: verschiedene Informationen werden parallel und über verschiedene Modalitäten verarbeitet.
- Simultan-redundante Interaktion: identische Informationen werden parallel und über verschiedene Modalitäten verarbeitet.

2.1.2.2 Potenzial und Einschränkungen multimodaler Mensch-Maschine-Schnittstellen

Durch die Integration von multimodalen Mensch-Maschine-Schnittstellen erreicht die Interaktion zwischen Mensch und Maschine eine völlig neue Qualität. Die Einbindung von Interaktionsparadigmen, die bereits aus der zwischenmenschlichen Kommunikation bekannt sind, führt zum einen zu einer intuitiveren Bedienbarkeit (Maybury und Wahlster, 1998; Oviatt et. al., 2001) und zum anderen zu effizienteren und fehlerrobusteren Dialogen (Althoff et. al., 2001; Cohen et. al., 2000; Net et. al., 2001; Oviatt, 2000a). Vor allem im Fahrzeug können multimodale Schnittstellen einen wichtigen Beitrag zur optimalen Aufgabebearbeitung leisten (Bengler, 2001).

Die Integration multimodaler Schnittstellen in serientaugliche Produkte ist jedoch mit zusätzlichen Aufwänden verbunden. In Anlehnung an Niedermaier (2003) müssen neben Kosten für zusätzliche Hardwarekomponenten (z.B. Kameras zur Gestenerkennung, Touchscreens, etc.) auch die Ausgaben für die Hardwarearchitektur (z.B. leistungsfähigere Prozessoren, mehr Arbeitsspeicher, etc.) erhöht werden. Ferner sind die Aufwände für die Entwicklung geeigneter Software sowie für die Spezifikation und das Design von multimodalen Dialogen im Vergleich zu herkömmlichen Systemen wesentlich höher. Dennoch sollte im Hinblick auf den steigenden Komplexitätsgrad von Userware die Integration multimodaler Schnittstellen weiter vorangetrieben werden, um deren Effizienz und Gebrauchstauglichkeit zu steigern.

2.1.3 Adaptive Mensch-Maschine-Schnittstellen und Benutzermodellierung

Wie bereits erläutert, leisten multimodale Mensch-Maschine-Systeme einen wichtigen Beitrag zur effizienteren Gestaltung von Interaktionen. Einen weiteren Baustein moderner Systeme bilden adaptive Komponenten. Häufig werden Begriffe wie Adaption, Adaptierbarkeit oder Personalisierung in wechselnden Bedeutungen angewendet. Allen gemein ist jedoch die Eigenschaft der Anpassung seitens der Maschine an den Menschen oder umgekehrt. Nach Stephanidis et. al. (1999) und Dix et. al. (2004) kann bei der Anpassung der Maschine an den Menschen grundlegend zwischen systeminitiiert und benutzerinitiiert Anpassung unterschieden werden. Während das Verhalten einer Mensch-Maschine-Schnittstelle bei systeminitiiert Modifikation automatisch durch das System angepasst wird, erfolgt die Modifikation bei benutzerinitiiert Anpassung durch den Benutzer selbst. Das System muss dazu geeignete Modifikationsmöglichkeiten bereitstellen. Findlater und McGrenere (2004) beschreiben als dritte Anpassungsart die gemischtinitiativ Anpassung, bei der Modifikationen sowohl durch den Benutzer als auch durch das System erfolgen können.

Im Rahmen dieser Arbeit wird der Begriff Adaption für systeminitiierte Anpassungen verwendet, die Begriffe Adaptierbarkeit bzw. Personalisierung für benutzerinitiierte Anpassungen. Biemans et. al. (2002) teilen Adaption in vier Kategorien ein: Adaption des Inhalts, Adaption der Funktionalität, Adaption der Präsentation bzw. Interaktion sowie Adaption qualitativer Art (z.B. Anpassung an geringe Bandbreiten).

Das große Potenzial der Adaption besteht darin, große Mengen an Information oder komplexe Strukturen derart zu reduzieren, dass ein Benutzer nicht überfordert wird und jederzeit effektiv und effizient mit einem technischen System arbeiten kann (Fischer, 2001). Ein entscheidender Aspekt dabei ist, bestimmte Eigenschaften, Fähigkeiten oder Präferenzen von Benutzern zu berücksichtigen. Diese spezifischen Merkmale müssen in einem Benutzermodell hinterlegt werden, welches nach Kobsa und Wahlster (1990) folgendermaßen definiert werden kann:

A user model is a knowledge source in a natural-language dialog system which contains explicit assumptions on all aspects of the user that may be relevant to the dialog behaviour of the system. These assumptions must be separable by the system from the rest of the system's knowledge.

Laut Kass und Finin (1988) können Benutzermodelle nach folgenden Dimensionen kategorisiert werden:

- Grad der Spezialisierung: Benutzermodelle können generisch oder individuell angelegt sein. Während generische Modelle eine homogene Grundgesamtheit

für verschiedene Nutzergruppen postulieren, enthalten individuelle Modelle ausschließlich spezifische Informationen über einen Nutzer.

- **Modifizierbarkeit:** Benutzermodelle können sich einerseits statisch verhalten, d.h. es findet keine Anpassung des Modells während einer Interaktion mit dem Nutzer statt. Bei dynamischen Modellen hingegen fließen Informationen über den Benutzer sofort in die Anpassung des Benutzermodells ein.
- **Zeitliche Gültigkeit:** mögliche Ausprägungen sind Kurzzeit- oder Langzeitmodelle. Während Kurzzeitmodelle z.B. während einer Interaktion zur Aufgabenbewältigung erstellt und anschließend wieder verworfen werden können, bleiben die gesammelten Informationen bei Langzeitmodellen als Wissen im System gespeichert und werden mit jeder weiteren Verwendung aktualisiert.
- **Anwendungsmethode:** beschreibende Benutzermodelle enthalten lediglich Informationen über den Benutzer in einer Datenbasis. Vorhersagende Benutzermodelle nutzen diese Informationen, um beispielsweise Verhalten und Reaktionen eines Benutzers zu simulieren und das System dementsprechend anzupassen.
- **Anzahl an Agenten:** ein System kann für die Verwendung durch eine Person ausgelegt sein oder für die Interaktion mit einer Person, die beispielsweise Interessen für eine weitere Person oder eine Gruppe vertritt. Ein Benutzermodell sollte wiederum selbst über entsprechende Modelle für diese Eventualitäten verfügen.
- **Anzahl an Modellen:** innerhalb eines Systems können ein einzelnes Benutzermodell oder mehrere Benutzermodelle dazu verwendet werden, eine Anpassung an den Anwender vorzunehmen.

Zusätzlich kann ein adaptives System klassifiziert werden, indem die Methodik zur Generierung eines Benutzermodells bewertet wird. Nach Kass und Finin (1988) kann die Erstellung eines Benutzermodells entweder implizit durch Analyse von Benutzerinteraktionen erfolgen oder explizit durch die Eingabe von Wissen oder Präferenzen durch den Benutzer.

2.2 Sprachdialogsysteme

2.2.1 Evolution von Sprachdialogsystemen

Die Entwicklung von ausgereiften Sprachdialogsystemen begann zunächst mit dem Einsatz als Auskunftssysteme im Telefoniebereich. Beispiele hierfür sind Flugauskunftssysteme (Peckham, 1991), Bahnauskunftssysteme (Aust et. al., 1995) oder auch Wetterauskunftssysteme (Zue et. al., 2000). Frühe Systeme waren statisch

strukturiert und ließen pro Dialogschritt nur eine begrenzte Anzahl an Eingaben zu. Jegliche Initiative ging dabei vom System aus.

„Für Kinoprogramm sagen Sie Kino, für Wetterbericht sagen Sie Wetter, ...“

Diese starren Dialoge der sog. Interactive Voice Response Systeme (IVR) wurden im Laufe der Zeit natürlicher und intuitiver gestaltet, indem gemischt-initiative Dialoge entworfen wurden. Im Vordergrund steht dabei die Erreichung eines bestimmten Ziels, wobei notwendige Information abwechselnd vom System und vom Benutzer angegeben werden. Gemischt-initiative Dialogsysteme sind daher in der Lage, dynamisch auf Eingaben zu reagieren und bieten großen Spielraum bei der Gestaltung natürlicher Dialoge (Hagen, 1999; Glass und Seneff, 2003; Zue und Glass, 2000).

Neben Auskunftssystemen halten komplexe Sprachdialogsysteme vor allem im Automobilbereich Einzug (Hagen et. al., 2004; Haller, 2003; Hamberger und Mauter, 2003; Heisterkamp, 2001). Die Dialoge zur Bedienung von Fahrerinformations- und Fahrerassistenzsystemen weisen dabei eine große Streuung bezüglich ihrer Komplexität auf. Das Spektrum reicht von kurzen Ein-Wort-Kommandos bis hin zu komplexen Dialogen wie der Zieleingabe im Navigationssystem. Die besondere Herausforderung im Fahrzeug besteht hauptsächlich darin, die Dialoge so zu gestalten, dass der Fahrer von seiner Primäraufgabe (Führung des Fahrzeugs) nur minimal abgelenkt wird.

2.2.2 Architektur eines Sprachdialogsystems

Ein Sprachdialogsystem ist nach Keller (2004) und McTear (2002) prinzipiell aus Komponenten zur Spracherkennung und zum Sprachverstehen, einer Komponente zur Dialogsteuerung und einer weiteren zur Sprachgenerierung und -synthese aufgebaut (siehe Abbildung 2.6). Im Folgenden werden die einzelnen Bestandteile eines Sprachdialogsystems kurz erläutert.

2.2.2.1 Spracherkennung und -verstehen

Die Extraktion einer Wortfolge aus einem Merkmalsvektor (Äußerung) erfolgt bei der automatischen Spracherkennung (ASE) durch die Bildung eines akustischen Modells und eines Sprachmodells. Ziel hierbei ist es, bei einer gegebenen Beobachtung $O = (o_1, \dots, o_T)$ mit T Merkmalsvektoren (Äußerungen) die tatsächlich geäußerte Wortfolge $W = (w_1, \dots, w_m)$ zu bestimmen. Dabei greift man auf die Wortfolge zurück, welche nach der Bayes'schen Formel die größte Wahrscheinlichkeit erhält:

$$P(W|O) = \frac{P(O|W) \cdot P(W)}{P(O)}$$

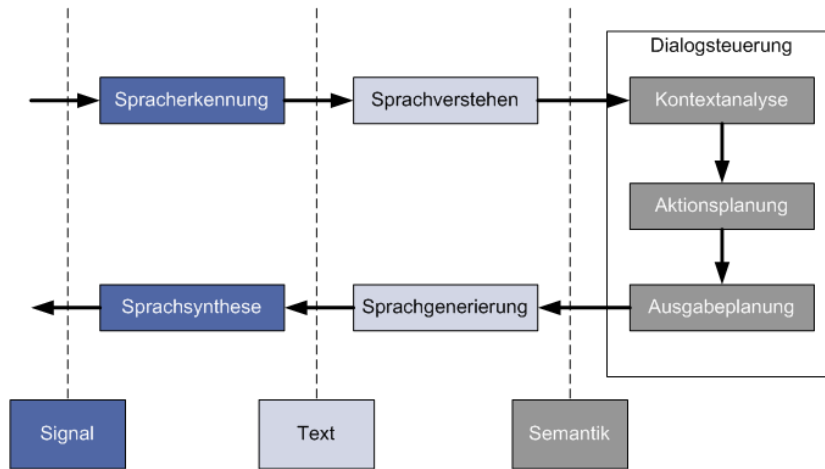


Abbildung 2.6: Architekturmodell für Sprachdialogsysteme nach Keller (2004) und McTear (2002)

Da der Nenner unabhängig von W ist, kann er an dieser Stelle vernachlässigt werden, womit lediglich das Produkt im Zähler maximiert werden muss:

$$P(W|O) = P(O|W) \cdot P(W) \rightarrow \max_W$$

Der Ausdruck $P(O|W)$ repräsentiert in diesem Fall das akustische Modell, welches mit Hilfe von Hidden-Markov Modellen (HMM), neuronalen Netzen oder Mischformen aus beiden Verfahren gebildet werden kann. Die größte Verbreitung finden jedoch HMM, mit deren Hilfe Übergangs- und Ausgabewahrscheinlichkeiten für bestimmte Merkmalsvektoren modelliert werden (siehe Abbildung 2.7)

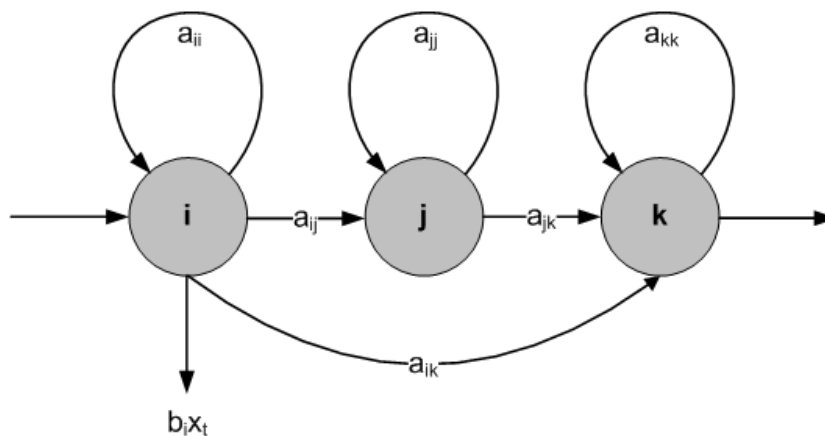


Abbildung 2.7: Grundprinzip von Hidden-Markov Modellen (Hohenner, 2004)

$P(W)$ bildet das Sprachmodell, in welchem Wissen über die Wahrscheinlichkeit von bestimmten Wortfolgen enthalten ist¹.

Als Sprachmodell kann entweder ein statistisches Modell (SLM) oder ein grammatikbasiertes Modell (GBM) angewendet werden. Bei einem SLM werden mögliche Wortfolgen durch stochastische Prozesse ermittelt, zumeist durch Training auf sehr großen Textkorpora. Im Gegensatz dazu erfolgt die Angabe möglicher Wortfolgen beim GBM manuell durch die Notierung in einer Grammatik.

Hinsichtlich der Leistungsfähigkeit spielen beide Sprachmodelltypen ihre Stärken in verschiedenen Anwendungsfällen aus. Erfolgen Spracheingaben durch Anwender innerhalb der festgelegten Grammatik, ist das GBM dem SLM gegenüber im Vorteil. Ist das jedoch nicht der Fall (z.B. falls ein Anwender das entsprechende Vokabular beim Erstkontakt nicht kennt), schneiden SLM besser ab (Gorrell, 2003; Hockey und Rayner, 2005).

2.2.2.2 Dialogsteuerung

Die Dialogsteuerung bildet das Herz eines SDS. Die Ergebnisse des Spracherkennungsprozesses werden hier interpretiert und unter Berücksichtigung verschiedener Faktoren wie der Diskurshistorie werden mögliche Interpretationsmöglichkeiten in einem bestimmten Kontext analysiert und bewertet.

In einem nachgelagerten Prozess werden die Interpretationen der Spracherkennung in konkrete Aktionen (z.B. Interaktionen mit Geräten) innerhalb eines Dialogzustands transformiert. Die Modellierung dieser Zustände erfolgt zumeist mittels einer Beschreibungssprache wie VoiceXML oder die Generic Dialogue Modeling Language (GDML) (Hanrieder und Hamerich, 2004; W3C, 2004), Beschreibungsstandards wie Speech Application Language Tags (SALT) oder die Extensible Multimodal Annotation Markup Language (EMMA) ermöglichen zudem die Modellierung multimodaler Dialoge (Cisco Systems Inc. et. al., 2002; W3C, 2005).

Nach der Aktionsplanung erfolgt die Planung der Ausgaben entsprechend der durchgeführten Aktionen. Je nachdem ob es sich um ein reines SDS oder ein multimodales System handelt, werden zudem Ausgaben in verschiedenen Modalitäten geplant.

2.2.2.3 Sprachgenerierung und -synthese

Systemausgaben (Prompts) können entweder durch vorgefertigte Sprachaufzeichnungen menschlicher Sprecher (Prerecorded Prompts) oder durch synthetische

¹Für eine erschöpfende Betrachtung der Grundlagen der ASE und von HMM sei an dieser Stelle auf Schukat-Talamazzini (1995) bzw. Carstensen et. al. (2004) verwiesen.

Sprachausgaben (Text-to-Speech) erfolgen. Bei der Sprachsynthese wird ein geschriebener Text mittels maschineller Verfahren in sprachliche Laute verwandelt. Ziel dabei ist es, die Ausgaben hinsichtlich ihrer Natürlichkeit (z.B. Prosodie) derart zu gestalten, dass diese nicht von menschlicher Sprache unterschieden werden können.

2.2.3 Vorteile sprachverarbeitender Systeme

Die Integration eines SDS im Fahrzeug birgt zwei entscheidende Pluspunkte in sich: zum einen ist Sprache eine natürliche und intuitive Kommunikationsform. Der Umgang mit Sprache muss nicht explizit erlernt werden, was bedeutet, dass die Sprachbedienung im Vergleich zu anderen Bedienkonzepten leichter zu erlernen ist (Wahlster, 1999). Zum anderen belegen Studien, dass die Verarbeitung akustischer Informationen im Fahrzeug Ablenkungspotenziale reduziert (Bengler, 1995). Im Vergleich zu visuellen oder haptischen Bediensystemen wirken sich nach Vollrath und Totzke (2000) akustische Bedienkonzepte nicht negativ auf die Fahrzeugführung (Längs- und Querverführung) aus und reduzieren zudem die körperliche und mentale Belastung (Erke et. al., 2001). Aus diesem Grund scheint die Implementierung eines SDS im Fahrzeug nicht nur sinnvoll sondern notwendig.

2.3 Gestaltungsaspekte beim Entwurf von Sprachdialogsystemen

Wie in Kapitel 2.1.1 dargestellt, müssen bei der Gestaltung von Useware und speziell von Sprachdialogsystemen die Eigenschaften eines Benutzers berücksichtigt werden. Unabhängig von einem festgelegten Anwenderkreis können, ausgehend von den Fähigkeiten und Limitationen des Menschen, Richtlinien dabei helfen, optimale Dialoge zwischen Mensch und Maschine zu entwerfen. Im Folgenden werden daher ergonomische und psychologische Aspekte näher beleuchtet, die für die Entwicklung insbesondere von Sprachdialogen von Bedeutung sind.

2.3.1 Softwareergonomische Richtlinien

Beim Entwurf von Softwaresystemen bietet die Europäische Norm EN ISO 9241-10 (Comité Européen de Normalisation, 1996) eine Orientierungshilfe für den Entwurf von gebrauchstauglichen Dialogen. Die Norm beinhaltet allgemeine Leitsätze für Arbeit an Bildschirmarbeitsplätzen, diese sind jedoch ausdrücklich unabhängig von einer bestimmten Dialogtechnik anwendbar. Eine Beschreibung der sieben Richtlinien ist in Abbildung 2.8 ersichtlich. Ergänzend dazu können die von Shneiderman (1997) aufgestellten Acht Goldenen Regeln des Interface Designs (siehe



Abbildung 2.8: Ergonomische Anforderungen für Büroarbeiten mit Bildschirmgeräten Teil 10: Grundsätze der Dialoggestaltung (EN ISO 9241-10)

Abbildung 2.9) herangezogen werden. Auch diese enthalten generelle Richtlinien, die beim Dialogdesign berücksichtigt werden sollten. Zusätzlich zu diesen Richtlinien existieren Empfehlungen, die speziell für Sprachdialoge entworfen wurden (Bernsen et. al., 1998; Krahmer et. al., 1999; Salmen, 2003). Zwar überschneiden sich viele der Richtlinien für Sprachdialoge mit den bereits dargestellten Empfehlungen, doch gründen sich die speziell für Sprachdialoge verfassten Richtlinien auf empirische Daten und lassen somit Rückschlüsse auf weitere wichtige Faktoren des Dialogdesigns zu. Ergänzend sollten daher die Aspekte in Betracht gezogen werden, die in Abbildung 2.10 aufgezeigt werden.

2.3.2 Kognitionspsychologische Aspekte

Neben softwareergonomischen Gesichtspunkten dürfen kognitionspsychologische Aspekte bei der Erstellung von Sprachdialogen nicht außer Acht gelassen werden. Nachfolgend werden daher grundlegende Eigenschaften des menschlichen Gedächtnisses und der Informationsverarbeitung erläutert.

2.3.2.1 Memorierungskapazität und mentale Belastung

Ein entscheidender Faktor für die Gestaltung effektiver Dialoge, ist eine genaue Kenntnis darüber, wie der Mensch Informationen verarbeitet und speichert. Da-

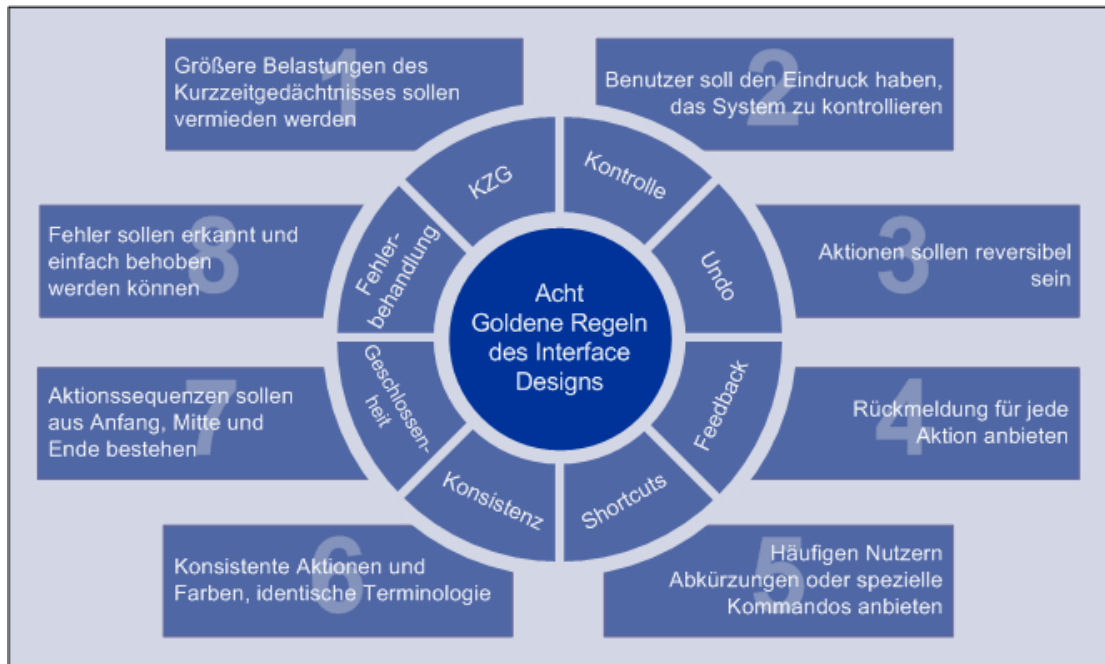


Abbildung 2.9: Acht Goldene Regeln des Interface Designs nach Shneiderman (1997)



Abbildung 2.10: Richtlinien für die Gestaltung von Sprachdialogen

bei ist vor allem von Interesse, wie viel Information ein durchschnittlicher Anwender im Kurzzeitgedächtnis (KZG) behalten kann. Auf dieser Grundlage muss die optimale Menge an Informationen bestimmt werden.

Welche Speicherkapazität im menschlichen KZG zur Verfügung steht, wurde erstmals von Miller (1956) untersucht. Diese begrenzte Speicherkapazität wird als

Gedächtnisspanne des KZG bezeichnet. Sie ist ein Maß für die maximale Anzahl an Informationseinheiten, die für kurze Zeit im Gedächtnis behalten werden kann. Die Anzahl ist von Mensch zu Mensch verschieden, die Versuche von Miller zeigten jedoch, dass das KZG im Durchschnitt Platz für $7(\pm 2)$ Informationseinheiten bietet (Miller, 1956; Baddeley, 1999; Zimbardo und Gerrig, 1999). Diese Einheiten können z.B. aus einzelnen Ziffern, Wörtern oder Satzteilen bestehen. Entscheidend ist folglich nicht die Anzahl der kleinsten Bestandteile (Items, z.B. Buchstaben eines Wortes), sondern allein die Anzahl der sog. Chunks.

Ein Chunk ist eine bedeutungstragende Informationseinheit. [...] Chunking ist der Prozess der Neuordnung (Rekodierung) einzelner Gedächtnisitems. Die Rekodierung kann durch Gruppierung auf der Basis von Ähnlichkeiten oder einem anderen Organisationsprinzip erfolgen. Sie kann aber auch in der Neukombination der Items zu größeren Mustern auf der Grundlage von Informationen bestehen, die im Langzeitgedächtnis gespeichert sind. (Zimbardo und Gerrig, 1999)

Die von Miller ermittelte Anzahl an Informationseinheiten basiert jedoch auf Werten, die unter Laborbedingungen gemessen wurden. In einer realen Umgebung muss von einer niedrigeren Zahl ausgegangen werden, da die mentale Belastung u.U. höher ist. Im Fahrzeug beispielsweise besteht die Primäraufgabe darin, das Fahrzeug im Straßenverkehr zu bewegen. Daneben werden für die Bedienung von FIS oder FAS ebenfalls Ressourcen benötigt, die insgesamt die Leistungsfähigkeit des KZG einschränken können (Wirth, 2002). Aus diesem Grund muss davon ausgegangen werden, dass speziell in der Domäne Fahrzeug die Gedächtnisspanne bei ca. $4(\pm 2)$ Informationseinheiten liegt.

2.3.2.2 Serielle Positionskurve

Die bisher erläuterten Erkenntnisse basieren auf der Grundlage, dass die Probanden eine bestimmte Abfolge von Items lernen und in der richtigen Reihenfolge wiedergeben mussten (serielle Wiedergabe). Dies ist im Umgang mit FIS oder FAS jedoch nicht nötig, die Items können in beliebiger Reihenfolge im Kurzzeitgedächtnis abgelegt bzw. wiedergegeben werden (freie Wiedergabe). Falls die freie Wiedergabe ohne Störung oder Unterbrechung erfolgt, kann bei der Wiedergabe der Items das Phänomen beobachtet werden, dass die ersten Items einer Liste gut reproduziert werden, ebenso die letzten Items. Man spricht hierbei vom Primacy- bzw. Recency Effekt (Baddeley, 1999; Zimbardo und Gerrig, 1999). Die Items in der Mitte der Liste können weniger gut wiedergegeben werden. Überträgt man diese Erkenntnisse in ein Diagramm, so lässt sich dafür eine charakteristische Kurve, die sog. Serielle Positionskurve ableiten (siehe Abbildung 2.11). Diese Eigenschaft ist speziell im Zusammenhang mit der Ausgabe von Informationen in Sprachdialogen von entscheidender Bedeutung. Fallen Sprachausgaben zu lang

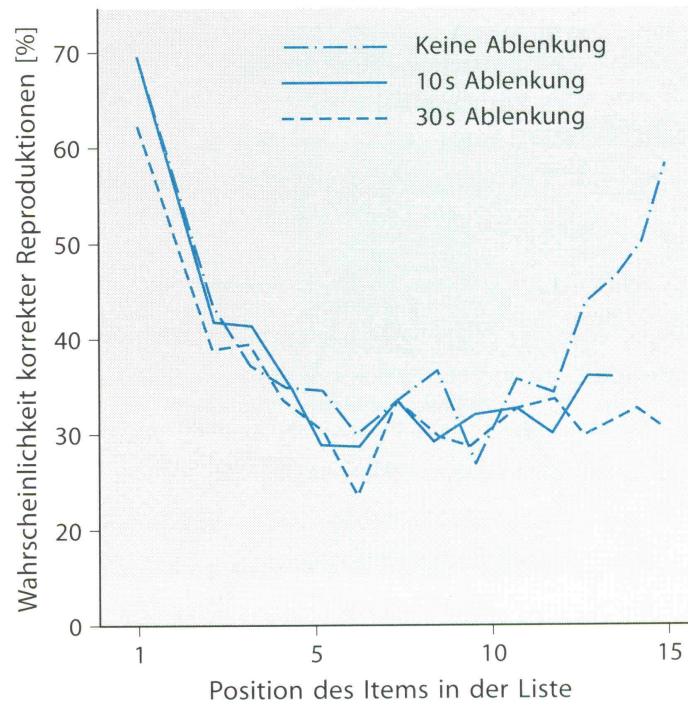


Abbildung 2.11: Serielle Positionscurve. Quelle: Zimbardo und Gerrig (1999)

aus bzw. enthalten diese zu viele Informationseinheiten, können diese von Fahrern nicht vollständig verarbeitet werden.

2.4 Fazit

Kapitel 2 vermittelt die Grundlagen aus dem Bereich der Mensch-Maschine-Interaktion. Einleitend werden die Begriffe Mensch-Maschine-System, Mensch-Maschine-Schnittstelle und Useware voneinander abgegrenzt. Gleichzeitig werden das IFIP-Modell und das Seeheim-Modell für Mensch-Maschine-Systeme vorgestellt. Von besonderer Bedeutung bei der Entwicklung von Mensch-Maschine-Systemen erweist sich die Erweiterung um multimodale und adaptive Komponenten, welche einen wesentlichen Beitrag zur Steigerung der Gebrauchstauglichkeit von Mensch-Maschine-Systemen im Allgemeinen und Sprachdialogsystemen im Speziellen leisten.

Das grundlegende Architekturmodell für SDS besteht aus Komponenten zur Spracherkennung, Dialogsteuerung und der Sprachsynthese. Im Rahmen dieser Arbeit kommt der Dialogsteuerung große Bedeutung zu, wobei beim Entwurf von Sprachdialogen verschiedene Gestaltungsaspekte berücksichtigt werden müssen.

Softwarergonomische Richtlinien wie etwa die EN ISO 9241-10 oder die Acht Goldenen Regeln des Interface Designs nach Shneiderman spielen beim Entwurf von Sprachdialogen eine ebenso große Bedeutung wie Aspekte aus dem Bereich der Kognitionspsychologie. Neben der Memorierungskapazität von 4 ± 2 Informationseinheiten speziell in der Fahrzeugdomäne spielt auch die Art und Weise des Abrufs von Informationen aus dem Gedächtnis eine Rolle. Der Primacy- und Recency-Effekt bewirkt eine gesteigerte Behaltensleistung bei Informationen, die zu Beginn bzw. am Ende einer Sprachausgabe stehen.

Kapitel 3

Stand der Technik im Bereich interaktiver Hilfesysteme

In Kapitel 3.1 erfolgt zunächst eine Definition des Begriffs Hilfesystem sowie die Einführung eines Schemas zur Klassifikation von Hilfesystemen. Kapitel 3.2 beinhaltet eine Übersicht über den Stand der Technik im Bereich interaktiver Hilfesysteme aus unterschiedlichen Anwendungsdomänen. Abschließend werden in Kapitel 3.2.8 existierende Hilfesysteme für SDS in Fahrzeugen vorgestellt.

3.1 Definition und Klassifikation von Hilfesystemen

Ausgehend von verschiedenen Vorstellungen darüber, was als Hilfesystem bezeichnet werden kann und welche Funktion dieses erfüllt, definiert sich ein Hilfesystem nach Bauer (1988) folgendermaßen:

Als **Hilfesystem** soll jedes Programm bezeichnet werden, das bei der Benutzung eines interaktiven Systems durch explizite Erklärungen hilft. Das System, für das die Hilfe gegeben wird, wird das **Zielsystem** genannt. Folgendes sind die wesentlichen Punkte dieser Definition:

1. Die Hilfe muß durch explizite Erklärungen erfolgen.
2. Der Problembereich, für den Hilfe gegeben wird, muß die interaktive Benutzung eines Computerprogramms betreffen.
3. Ein Hilfesystem ist selbst ein Computerprogramm.

Basierend auf dem jeweiligen Informationsbedarf können nach Krause (1988) zwei verschiedene Arten von Hilfesystemen abgeleitet werden. Auf der einen Seite existieren tutorielle Systeme, welche einen generellen Informationsbedarf abdecken und „[...] den Benutzer beim systematischen Wissenserwerb unterstützen“ (Mittermaier, 1995). Auf der anderen Seite können Auskunftssysteme angeführt wer-

den, welche einen spezifischen Informationsbedarf bedienen und der unmittelbaren Lösung einer Aufgabe dienen.

Laut Bauer (1988) kann eine weitergehende Klassifikation nach folgenden Kriterien erfolgen:

- **Initiative (aktiv vs. passiv):** Aktive Hilfesysteme besitzen die Fähigkeit, dem Benutzer selbstständig Hilfe anzubieten. Über die Protokollierung von Benutzerinteraktionen können beispielsweise suboptimale Bedienabfolgen identifiziert und dementsprechend Hilfestellungen geleistet werden. Die Initiative liegt hierbei beim System selbst. Im Gegensatz dazu wird passive Hilfe stets vom Benutzer initiiert und dient einer gezielten Suche nach Information.
- **Kontextbezug (statisch vs. dynamisch):** Statische Hilfe ist kontextunabhängig und bietet dem Benutzer zu jedem Zeitpunkt die gleichen Inhalte an. Als Beispiel kann das UNIX Hilfesystem MAN angeführt werden, welches zu jedem Kommando die entsprechende Syntax, Parameter und Beispiele ausgibt. Dynamische Hilfesysteme dagegen berücksichtigen den aktuellen Kontext, in dem die Hilfe angefordert wird. Ein solches System ist im iDrive SDS integriert. So erzeugt das Hilfefunktion „Optionen“ bezogen auf den Kontext (z.B. Radio oder Navigation) eine entsprechende Auflistung der möglichen Sprachkommandos.
- **Individualität (einheitlich vs. individuell):** Ein Großteil der existierenden Hilfesysteme bietet eine einheitliche Hilfe an, d.h. jeder Benutzer erhält unabhängig von Präferenzen oder bereits vorhandenem Wissen die gleiche Hilfe. Um individuelle Hilfe anbieten zu können, muss spezifisches Wissen des Benutzers in einem Benutzermodell abgelegt werden. Basierend auf diesem Benutzermodell kann für jeden Benutzer eine optimierte Hilfe angeboten werden.
- **Hilfezeitpunkt (synchron vs. asynchron):** Bei passiven Hilfesystemen erfolgt die Hilfeausgabe meist synchron, d.h. die Ausgabe folgt unmittelbar auf eine Hilfeanfrage. Synchrone Hilfeausgaben bedienen einen konkreten Informationsbedarf des Benutzers. Die asynchrone Ausgabe macht daher nur bei aktiven Hilfesystemen Sinn, da unter Umständen kein konkreter Handlungs- oder Informationsbedarf vorliegt. So kann beispielsweise durch eine zeitlich verzögerte Ausgabe eine unnötige Belastung des Benutzers vermieden werden.
- **Integrationsgrad (anwendungsabhängig vs. anwendungsunabhängig):** Anwendungsabhängige Hilfesysteme sind vollständig in das Zielsystem integriert und können nicht separat betrieben werden. Als Beispiel kann abermals das Hilfesystem des iDrive SDS angeführt werden. Im Gegensatz dazu ist anwendungsunabhängige Hilfe vom Zielsystem getrennt. Diese Hilfesysteme können in verschiedenen Anwendungsbereichen eingesetzt werden,

so z.B. das Hilfesystem HTMLHelp des Microsoft Windows Betriebssystems, welches von diversen Windowsapplikationen angesprochen werden kann.

3.2 Konzepte interaktiver Hilfesysteme

Hilfesysteme sollen einen Benutzer bei der Durchführung einer Aufgabe unterstützen. Das Spektrum von Hilfesystemen kann von einem Glossar oder einer Sammlung häufig gestellter Fragen (Frequently Asked Questions FAQ) bis zu aktiven intelligenten Hilfesystemen reichen.

Die Betrachtung tutorieller Systeme wird an dieser Stelle ausgeklammert. Tutorielle Systeme vermitteln zumeist breiteres Wissen über Gegenstandsbereiche, als für die Lösung einer konkreten Aufgabe nötig ist. Insbesondere im Fahrzeug müssen Hilfesysteme Informationen so effektiv und effizient als möglich anbieten, um potenziell störende Einflüsse auf die Fahrzeugführung zu minimieren. Ferner sind die pädagogischen und didaktischen Anforderungen an die Entwicklung tutorieller Systeme sehr komplex und können auf Grund ihres Umfangs nicht Gegenstand dieser Arbeit sein.

Entsprechend der Zielsetzung dieser Arbeit, ein Hilfesystem für SDS in automotiven Umgebungen zu entwickeln, beschränkt sich die Standortbestimmung auf Hilfesysteme, die konzeptuell für den Einsatz im Fahrzeug geeignet sind.

3.2.1 COMFOHELP

Krause et. al. (1993) stellen mit COMFOHELP ein Hilfesystem für das Textverarbeitungsprogramm COMFOTEX vor. Die Kernkomponente des Hilfesystems bildet ein Planerkenner, der mögliche Dialogziele des Benutzers ermittelt und dementsprechend aktiv oder passiv Hilfestellung anbieten kann. COMFOHELP verwendet ähnlich wie andere Planerkennungssysteme (z.B. SINIX Consultant, Hecking (2000)) eine Planbibliothek als Grundlage für die Analyse von Bediensequenzen. Die Aktionen eines Benutzers werden beobachtet (Keyhole Recognition, vgl. Ardissono (1996)) und mit einer Planbibliothek abgeglichen. Die Planbibliothek umfasst Wissen über verschiedene Dialogziele in Form von vordefinierten Plänen folgender Ausprägung:

- Optimale Pläne: Liste mit optimalen Aktionssequenzen für jedes Dialogziel
- Suboptimale Pläne: Liste mit Aktionssequenzen, die ein Dialogziel mit mehr Interaktionsschritten als nötig beschreiben
- Möglicherweise falsche Pläne: Aktionssequenzen, die kein Dialogziel spezifizieren, in einem anderen Kontext jedoch sinnvoll sein können

- Falsche Pläne: Liste mit Aktionssequenzen, mit denen ein bestimmtes Dialogziel nicht erreicht werden kann

Die Notation der Pläne für ein Dialogziel erfolgt tabellarisch, wobei jeder Plan einzeln bewertet wird. Quelltext 3.1 zeigt eine Liste alternativer Pläne für das Dialogziel „Cursor an den Anfang des Dokuments bewegen“. Die Tastenkombination $\langle \text{Strg} + \text{Pos1} \rangle$ wird als optimaler Plan o eingestuft, mit a werden beinahe gleichwertige Alternativen über das Menü „Bearbeiten“ bzw. das direktmanipulative Positionieren des Cursor C am Anfang des Dokuments gekennzeichnet. Vorheriges Scrollen oder die Positionierung des Cursors durch Verwendung der Pfeiltasten in begrenztem Umfang ($n < x$) bewirkt eine Kategorisierung als leicht suboptimal (ls), wobei die extensive Verwendung der Pfeiltasten ($n > x$) suboptimal ist (s). Das Wissen des Systems umfasst neben den Plänen zusätzlich sog. Durch-

C_DokAnfang	o	$\langle \text{Strg} + \text{Pos1} \rangle$
	a	BEARBEITEN/Gehe_zu; Anfang; OK
	a	C [Zielpos=DokAnfang]
	ls	Scrollen; C [Zielpos=DokAnfang]
	ls	$n < \text{CTasten} \rangle$ [Zielpos=DokAnfang; $n < x$]
	s	$n < \text{CTasten} \rangle$ [Zielpos=DokAnfang; $n > x$]

Quelltext 3.1: COMFOHELP Pläne für das Dialogziel „Cursor an den Anfang des Dokuments bewegen“

laufkommandos. Diese Kommandos können zwar Teil einer Aktionssequenz sein, beeinflussen die Zuordnung zu einem Plan jedoch nicht.

Der Abgleich der beobachteten Sequenzen mit der Planbibliothek erfolgt durch einen Parser. Sobald ein Plan erkannt wird, reagiert das Hilfesystem differenziert auf verschiedene Plantypen:

- Optimaler Plan: der Plan wird aus der Planbibliothek gelöscht, da angenommen wird, dass dem Benutzer die optimale Bediensequenz bekannt ist.
- Suboptimaler Plan: nach der dritten Erkennung des identischen Plans gibt das System eine Hilfenachricht aus.
- Möglicherweise falscher Plan:
 - Wird unmittelbar auf die erkannte suboptimale Sequenz die passive Hilfe aufgerufen, gibt das System eine Hilfenachricht zu dieser Sequenz aus.
 - Existieren mehrere mögliche suboptimale Pläne, fragt das System explizit nach dem gewünschten Hilfethema („Wollten Sie Text kopieren oder ausschneiden?“).
- Falscher Plan: nach dreimaliger Erkennung initiiert das System einen Klärungsdialog.

3.2.2 PHI

PHI ist ein Hilfesystem, das für die UNIX-Domäne entwickelt wurde und prototypisch für die *mailman* Anwendung implementiert wurde (Bauer et. al., 1993). Ebenfalls auf Planerkennung basierend unterscheidet sich PHI von COMFOHELP zum einen beim Aufbau der Planbibliothek und zum anderen in der Analyse der Pläne. Der Aufbau der expliziten Planbibliothek bei PHI erfolgt durch zwei sog. *cross-talk modes* zwischen der Plangenerierungskomponente und dem Planerkenner. Der Plangenerator enthält zunächst nur abstrakte Pläne zu Dialogzielen. Im ersten cross-talk mode übergibt der Plangenerator die abstrakten Pläne an den Planerkenner, der die Pläne mit konkreten Ausprägungen instanziiert. Falls die Analyse des Benutzerverhaltens eine suboptimale Aktionssequenz für ein Dialogziel ergibt, erfolgt im zweiten cross-talk mode die Berechnung eines optimalen Plans für dieses Ziel durch den Plangenerator.

Die Repräsentation der Pläne und Aktionssequenzen erfolgt auf Basis der *logical language for planning* (LLP) (Biundo et. al., 1992). LLP ist eine intervallbasierte modale Temporallogik, die neben Konzepten aus der Prädikatenlogik die modalen Operatoren „ \circ “ (next), „ \diamond “ (sometimes) und „ \square “ (always) sowie den binären Operator „ $;$ “ (chop) bereitstellt, der die Modellierung zeitlicher Abfolgen ermöglicht. Pläne haben in LLP folgende Struktur:

$$\mathbf{if} \ df((x, M) = T \ \mathbf{then} \ EX(undelete(x, M)) ; EX(type(x, M))$$

Dieser Plan beschreibt das Lesen einer Mail x aus der Mailbox M mit der Aktion $type$. Die Mailbox muss dazu bereits geöffnet sein und die Mail darf nicht gelöscht sein ($df(x, M)$). Falls dies doch der Fall sein sollte ($df(x, M) = T$), wird die Mail wieder hergestellt ($undelete(x, M)$) und kann anschließend gelesen werden ($EX(type(x, M))$).

Die Spezifikation von Plänen basiert auf LLP-Formeln mit folgendem Muster:

$$[Vorbedingung \wedge \mathbf{Plan}] \rightarrow \mathbf{Ziel}$$

Die Planspezifikation „Inhalt der Mailbox M auf dem Display anzeigen und anschließend Mail x lesen“ hat demnach folgendes Aussehen:

$$of(M) = T \wedge \mathbf{Plan} \rightarrow \diamond[displ = headers(M) \wedge \diamond[rf(x, M) = T]]$$

Aufgabe des Plangenerators ist, für die Variable *Plan* ein Axiom zu finden, welches die Aussage wahr werden lässt; in diesem Fall:

$$EX(header(M)) ; \mathbf{if} \ df(x, M) = T \ \mathbf{then} \ EX(undelete(x, M)) ; EX(type(x, M))$$

Entsprechend des ermittelten Axioms, kann nun aktive Hilfe zur Erreichung eines Ziels ausgegeben werden. Ergeben sich in der Plangenerierungsphase mehrere gültige Axiome, erfolgt die Auswahl eines Axioms auf Basis probabilistischer Selektionsverfahren.

3.2.3 Lumière (Microsoft Office Assistent)

Horvitz et. al. (1998) präsentieren im Rahmen des Lumière-Projekts ein Hilfesystem für Microsoft Excel, das mögliche Ziele eines Benutzers ermittelt und dafür aktive oder passive Hilfe zur Verfügung stellen kann. Grundlage hierbei sind jedoch keine Planerkennungsalgorithmen, sondern eine Temporallogik angewendet auf Bayes'sche Netze. Die Bestimmung möglicher Ziele ($Goal_{t_0}$) zum Zeitpunkt t_0 erfolgt durch Beobachtungen vorangegangener Aktionen (E_i) des Benutzers. Das Lumière-System bildet ein temporales Modell der Aktionen und berechnet die Wahrscheinlichkeitswerte $p(E_{i,t}|Goal_{t_0})$ bestimmter Ziele unter der Bedingung verschiedener Aktionsfolgen.

Der Wertebereich möglicher Wahrscheinlichkeiten liegt zwischen der Wahrscheinlichkeit für eine unmittelbar beobachtete Aktion $p(E_{i,t_0}|Goal_{t_0})$ und einer Aktion außerhalb der Beobachtungsfolge $p(E = false|Goal_{t_0})$. Die Gültigkeit der einzelnen Wahrscheinlichkeiten wird durch sog. *evidential horizons* (EH) begrenzt. EH markieren Schwellenwerte, an denen die Wahrscheinlichkeiten für ein Ziel neu berechnet werden. Diese Schwellenwerte werden von Experten definiert und können Zeitspannen oder eine festgelgte Anzahl an Interaktionsschritten sein. Wird ein EH erreicht, definieren Zerfallsfunktionen (z.B. lineare oder exponentielle Funktionen) die weitere Bewertung der Wahrscheinlichkeit eines Ziels. Auch die Zerfallsfunktionen werden von Experten definiert.

Die Anbindung des Hilfesystems an Microsoft Excel erfolgt durch die Lumière Events Language (LEL). Atomare Aktionen wie das Klicken mit der Maus auf einen Menüeintrag werden protokolliert und analysiert. Zur Verfügung stehen beispielsweise folgende Methoden:

- $Rate(x_i, t)$: Häufigkeit, mit der eine atomare Aktion x_i in t Sekunden oder Interaktionen verwendet wird
- $OneOf(\{x_1, \dots, x_n\}, t)$: mindestens eine Aktion einer definierten Aktionssequenz wird in der Zeit t verwendet
- $All(\{x_1, \dots, x_n\}, t)$: alle Aktionen einer definierten Aktionssequenz werden in der Zeit t verwendet
- $Seq(\{x_1, \dots, x_n\}, t)$: eine Sequenz von festgelegten Aktionen wird in der Zeit t verwendet, andere Aktionen können z.T. in der Sequenz enthalten sein
- $TightSeq(\{x_1, \dots, x_n\}, t)$: eine Sequenz von festgelegten Aktionen ohne andere Aktionen wird in der Zeit t verwendet

- *Dwell(t)*: im Zeitraum t werden keine Aktionen beobachtet

Auf diese Weise können Rückschlüsse über das Benutzerverhalten gezogen werden, z.B. wie viele Sekunden ohne Aktivität nach einer bestimmten Aktion verstreichen. Informationen dieser Art werden in einem persistenten User Model in der Windows Registry gespeichert. Der Aufbau des User Models kann durch drei Strategien erfolgen:

- Pulsed Strategy: die Analyse von Aktionen und Inferenzoperationen finden in festen Zeitintervallen statt
- Event Driven: Analyse und Inferenz werden durch eigens markierte Schlüsselaktionen ausgelöst
- Augmented Pulsed Strategy: Analyse und Inferenz finden statt, wenn das System keine große Systemauslastung feststellt

Ermittelt das Hilfesystem während der Systembedienung ein potenzielles Dialogziel, wird eine systeminitiierte Hilfe ausgegeben, welche die wahrscheinlichsten Hilfethemen abdeckt. Der Benutzer kann mit Hilfe einer sog. „Volume Control“ bestimmen, wie stark das Hilfesystem sich in die Dialoge einbinden soll. Konzepte und Themengebiete, die bereits zu einem früheren Zeitpunkt behandelt wurden, durch den Benutzer jedoch mehrmals ignoriert wurden, werden aus der Hilfe ausgeschlossen, bis eine signifikante Änderung des Benutzerverhaltens auftritt.

Eine funktional stark reduzierte Variante dieses Konzepts findet seine Anwendung im Microsoft Office Assistenten (Horvitz et. al., 1998). Der Assistent leitet anhand der durchgeführten Interaktionen eine Liste mit möglichen Zielen und resultierenden Hilfethemen ab, welche dem Benutzer angeboten wird. Ausführliche Hilfeinformationen können durch die Auswahl eines Hilfethemas abgerufen werden.

3.2.4 Kyoto Sightseeing Guide

Fukubayashi et. al. (2006) stellen ein aktives Hilfesystem für ein sprachgesteuertes Auskunftssystem vor, welches das Wissen eines Benutzers über das System bei der Ausgabe der Hilfe berücksichtigt. Sowohl die Struktur des Systems als auch das Wissen des Benutzers wird durch einen *Domain Concept Tree* (DCT) bestehend aus Knoten (Konzepte) und Kanten (Abhängigkeiten) modelliert. Der hierarchische Konzeptbaum besteht aus vier Ebenen:

- Systemebene: repräsentiert das gesamte System
- Funktionsebene: beinhaltet abstrakte Funktionen des System, z.B. Informationen zu Tempeln erfragen
- Elementebene: besteht aus abstrakten Elementen, die konkrete Äußerungen enthalten können, z.B. die Bezeichnung von Tempeln

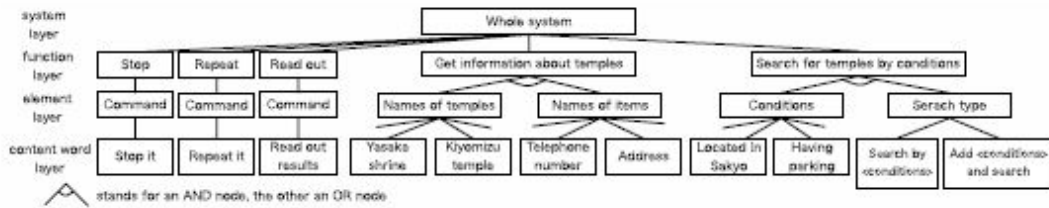


Abbildung 3.1: Domain Concept Tree. Quelle: Fukubayashi et. al. (2006)

- Wortebene: beinhaltet das Vokabular, das in den Elementen enthalten sein kann, z.B. „Kiyomizu Tempel“

Es existieren zwei Arten von Knoten; AND Knoten kommen nur in der Funktionsebene vor und legen fest, welche Elemente in einer vollständigen Anfrage des Benutzers enthalten sein müssen. OR Knoten können in allen Ebenen vorkommen und bezeichnen alternativ anwendbare Konzepte.

Jeder Knoten verfügt über einen zugewiesenen Wert p , welcher ausdrückt, wie gut ein Benutzer ein bestimmtes Konzept kennt. Anfänglich erhält jeder Knoten den Wert $p = 0,5$. Werte über $0,5$ bedeuten, dass ein Konzept bekannt ist, Werte kleiner $0,5$ kennzeichnen einen Hilfebedarf zu diesem Konzept. Die Neuberechnung von p für jeden Knoten n erfolgt nach einer Benutzeräußerung oder einer Hilfeausgabe durch eine Funktion $renew(n, p)$. Für die Veränderung von p existieren verschiedene Methoden:

- $renew(n, p_u^+)$: nach einer Benutzeräußerung wird der Wert p eines Knotens n um p_u erhöht.
- $renew(n, p_u^-)$: nach einer Benutzeräußerung wird der Wert p eines Knotens n um p_u erniedrigt.
- $renew(n, p_h^+)$: nach einer Hilfeausgabe wird der Wert p eines Knotens n um p_h erhöht.

Die Anwendung der einzelnen Methoden kann anhand einer schematischen Darstellung des DCT (siehe Abbildung 3.2) deutlich gemacht werden.

Sobald eine Benutzeräußerung erfolgt, wird das Erkennerergebnis analysiert und mit der Wortebene abgeglichen. Enthält das Erkennerergebnis beispielsweise die Vokabel E , wird dessen Wert mittels $renew(E, p_u^+)$ um p_u^+ erhöht, da Konzept E dem Benutzer bekannt ist. Anschließend wird geprüft, ob der Vorgängerknoten (C) ein AND Knoten ist, und ob dessen Bedingung erfüllt ist. In diesem Fall ist C ein AND Knoten, die Bedingung ist jedoch nicht erfüllt, da D nicht in der Äußerung enthalten war. Der Wert des Knotens C wird daher mittels $renew(C, p_u^-)$ um $\lambda^1 \cdot p_u^-$ erniedrigt. Der Operator λ beschreibt die Gewichtung von p_u . Diese nimmt mit zunehmender Entfernung von ursprünglich betrachteten Knoten E

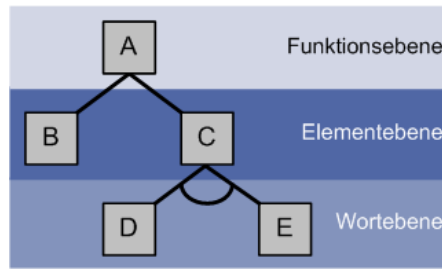


Abbildung 3.2: Schematische Darstellung des DCT

ab $(\lambda^0, \lambda^1, \lambda^2, \dots)$, da nicht sicher darauf geschlossen werden kann, dass das Wissen über ein Konzept E übergeordnete Konzepte (C, A) in gleicher Form mit einschließt. Der Knoten A erhält somit den Wert $p_A = \lambda^2 \cdot p_u^-$.

Nach der Ausgabe einer Hilfenachricht zu Konzept A , erhält dieser Knoten den Wert $p_A = p_h^+$. Eventuell enthaltene konkrete Konzepte (Beispiele) aus der Wortebene (z.B. D und E) werden ebenfalls neu gewichtet. Die Knoten D und E erhalten jeweils den neuen Wert $p_{D,E} = \frac{p_h^+}{N}$, wobei N die Anzahl der Beispiele repräsentiert. Diese Maßnahme soll die Annahme widerspiegeln, dass mehrere Beispiele schwerer zu memorieren sind als eines.

Die Ausgabe der aktiven Hilfe erfolgt entweder bei unvollständigen Eingaben, d.h. eine AND Bedingung wurde nicht erfüllt oder nach einer Zeitspanne ohne Äußerung. Startknoten für die Berechnung einer Hilfenachricht ist der Knoten, für den zuletzt $renew(n, p_u^-)$ ausgeführt wurde. Der für die Hilfeausgabe mögliche Konzeptknoten muss folgende Bedingungen erfüllen:

1. Er ist Nachfolger des Startknotens oder der Startknoten selbst.
2. Der Wert p des Knotens muss unter 0,5 liegen.
3. Er liegt in der dem Startknoten nächsten Ebene, die Kriterien 1 und 2 erfüllt.

Liefert die Analyse der Knoten mehrere mögliche Konzepte, wird das mit dem kleinsten Wert p ausgewählt. Existieren mehrere Knoten mit gleich niedrigen Werten, wird ein zufälliger für die Hilfe ausgewählt. Abhängig von der Ebene, in der sich ein Konzept befindet, existieren für jede Ebene Vorlagen für Hilfetexte, in die die Beschreibung der möglichen Konzepte eingefügt wird (vgl. Beispieldialog 3.1).

3.2.5 SmartAidè

Ramachandran und Young (2005) entwickeln mit SmartAidè ein Hilfesystem für graphische Benutzeroberflächen, welches dem Benutzer eine detaillierte Anleitung zu verschiedenen Hilfethemen einer MS Windows Anwendung (z.B. iTunes) bietet.

User:	Tell me about Kiyomizu Temple .
System:	I could not understand what you want to know about <i>Kiyomizu Temple</i> . To <i>get information about temples</i> , ask me a question with both <i>a name of a temple</i> and <i>an item of the temple</i> .
User:	Tell me the history of Kiyomizu Temple .
System:	I could not understand what you want to know about <i>Kiyomizu Temple</i> . You can <i>get information about about temples</i> , for example, by saying <i>summary, adress, telephone number, as a name of items</i> .

Dialog 3.1: Kyoto Sightseeing Guide: fett dargestellte Dialogteile sind die erkannten Wörter des Spracherkenners, kursiv dargestellte Dialogteile wurden vom Hilfesystem in die Antwortvorlagen für die verschiedenen Ebenen des Auskunftssystems eingefügt (vgl. Abbildung 3.1).

Sobald eine Applikation gestartet wird, öffnet sich auch ein SmartAidè Dialog. Das Hilfesystem beobachtet alle Zustände von Objekten innerhalb der Applikation sowie den Zustand der Applikation selbst (minimiert, maximiert).

Jede ausgeführte Aktion wird mit einer Aktionsbibliothek abgeglichen, in welcher verschiedene Aktionssequenzen hinterlegt sind. Jeder Eintrag in der Aktionsbibliothek enthält Informationen über notwendige Vorbedingungen und mögliche Nachfolgebedingungen. Durch den Abgleich der beobachteten Aktionen mit der Aktionsbibliothek versucht SmartAidè, mögliche Ziele des Benutzers zu identifizieren und eine optimale Hilfe zu generieren.

Der Aufruf der Hilfe erfolgt durch die Auswahl eines Hilfethemas aus einer Liste möglicher Hilfethemen, welche in Abhängigkeit von den Interaktionen des Benutzers aktualisiert wird. Wählt der Benutzer ein Hilfethema aus, wird ein entsprechender Plan zum Erreichen dieses Ziels erzeugt. Dabei werden Informationen über den aktuellen Kontext und zum Zielzustand ausgewertet, um eine mögliche Aktionssequenz zwischen Ist- und Zielzustand zu erzeugen. Die Auswertung von Informationen über die Reihenfolge von Aktionen und Vorbedingungen, die eventuell bereits erfüllt sind, liefert einen an den momentanen Kontext angepassten Plan.

Die einzelnen Aktionen dieses Plans werden in XML kodiert und an einen Execution Manager (EM) gesendet, der einen gerichteten zyklischen Graphen daraus erzeugt. Die Knoten des Graphen repräsentieren Aktionen, die Kanten bilden kausale und zeitliche Abhängigkeiten ab. Der EM entfernt ggf. überflüssige Knoten und transferiert jede Aktionen und deren Parameter des resultierenden Plans in ein XML Beschreibungsformat, das konkrete Funktionsaufrufe für die Zielapplikation beschreibt. Dieses XML Schema wird an das SmartAidè Hilfemodul gesendet, welches die Funktionsaufrufe ausführt und auf diese Weise Schritt für Schritt Objekte einer Applikation verändert. Gleichzeitig werden zeitlich synchron entsprechende textuelle Hilfeausgaben produziert. Auf diese Weise erhält der Benutzer

sowohl eine Demonstration als auch eine Erläuterung der notwendigen Aktionen zum gewünschten Zielzustand.

3.2.6 Targeted Help

Unter dem Begriff Targeted Help (TH) stellen Hockey et. al. (2003) ein Hilfekonzzept für SDS vor, welches die Domäne von Out-Of-Vocabulary (OOV) Äußerungen abdeckt. OOV bedeutet, dass eine Benutzereingabe Vokabular enthält, das nicht in einer Grammatik spezifiziert wurde und somit vom SDS nicht verarbeitet werden kann. In der Regel reagiert eine SDS auf diese Dialogsituationen, indem eine Aufforderung zur nochmaligen Eingabe eines Kommandos ausgegeben wird. Unter Verwendung von TH erfolgt jedoch eine Analyse der OOV Äußerung mit Hilfe eines zweiten Spracherkenners mit statistischem Sprachmodell (vgl. Kapitel 2.2.2.1).

Das TH Modul ist getrennt vom regulären SDS und beinhaltet einen SLM Spracherkenner, einen Targeted Help Activator zur Analyse der Erkennerergebnisse des ersten und zweiten Spracherkenners und einen Targeted Help Agent, der die Hilfe ausgibt (vgl. Abbildung 3.3). Der Targeted Help Activator steuert die Hilfe

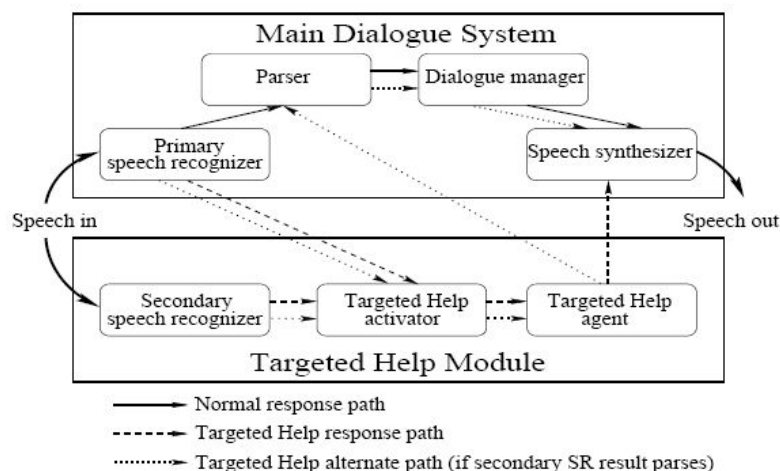


Abbildung 3.3: Aufbau des Targeted Help Moduls. Quelle: Hockey et. al. (2003)

unter Berücksichtigung verschiedener Erkennerergebnisse:

- Liefern sowohl der erste grammatikbasierte Spracherkenner (GB) als auch der zweite Spracherkenner (SLM) ein eindeutiges Erkennerergebnis, bleibt TH deaktiviert. Der Dialog wird mit dem Resultat des GB Erkenners fortgesetzt.

- Ist das Ergebnis des GB Erkenners eindeutig, das des SLM Erkenners jedoch nicht, bleibt die Hilfe ebenfalls deaktiviert.
- Liefert der GB Erkennen kein Ergebnis, der SLM Erkennen jedoch schon, wird zunächst versucht, das Ergebnis des SLM Erkenners zu parsen. Falls das Parsing erfolgreich verläuft, wird das Ergebnis an den GB Erkennen gesendet und dient dort der Dialogfortsetzung.
- Erzeugt der GB Erkennen ein schlechtes, der SLM Erkennen ein gutes Ergebnis und schlägt das Parsing des SLM Ergebnisses fehl, wird der Targeted Help Agent aktiviert.
- Liefern sowohl GB als auch SLM Erkennen ein schlechtes Erkennenresultat, kann keine Interpretation der Eingabe erfolgen und die Standardfehlernachricht wird ausgegeben.

Die Aktivierung des Targeted Help Agents setzt einen Analyseprozess in Gang. Zunächst wird geprüft, was das Hilfesystem verstanden hat. Die Diagnose der Eingabe kann drei verschiedene Ausprägungen annehmen:

- Endpoint Fehler: der Anfang oder das Ende einer Äußerung wurde abgeschnitten. Die Identifizierung dieses Fehlers basiert auf einem Abgleich des SLM Erkennenresultates mit der Grammatik des GB Erkenners.
- Unbekanntes Vokabular: das bei der Eingabe verwendete Vokabular ist nicht in der Grammatik des GB Erkenners, aber im SLM Vokabular enthalten. Das Hilfesystem erzeugt dann eine Ausgabe der Form „The system doesn’t understand the word X“.
- Subkategorisierungsfehler: ein bestimmtes Wort ist zwar in der Grammatik enthalten, die Art und Weise wie es verwendet wurde entspricht jedoch keiner gültigen Regel. So kann z.B. die Äußerung „Fly between the hospital and the school“ den gültigen Befehl „fly“ enthalten, die Parameter „between the hospital and the school“ jedoch sind in der Form ungültig. Die Hilfeausgabe würde in diesem Fall lauten: „The system heard *fly between the hospital and the school*, unfortunately it doesn’t understand *fly* when used with the words *between the hospital and the school*. You could try saying *fly to the hospital*“. Die Identifizierung dieses Fehlertyps erfolgt ebenfalls durch den Abgleich des Erkennenresultates mit der Grammatik.

Die Ausgabe der Hilfetexte erfolgt grundsätzlich mit gültigem Vokabular aus der Grammatik, das vom Benutzer bereits richtig angewendet wurde. Auf diese Weise erhält der Benutzer auch eine implizite positive Rückmeldung über die eigenen Äußerungen.

3.2.7 MATCH: Multimodal Access to City Help

Hastie et. al. (2002) entwarfen ein multimodales Hilfesystem für das Auskunftssystem MATCH (Multimodal Access to City Help). Die Hilfedomäne umfasst drei Bereiche:

- Informationen zu Restaurants
- Auskunft über U-Bahn-Verbindungen
- Hilfe zur Kartenbedienung

Der Aufruf der Hilfe kann durch das Sprachkommando „Hilfe“, das Schreiben des Worts „Hilfe“ mit einem Stift auf dem Kartenbildschirm oder durch Antippen eines Hilfebuttons mit dem Stift erfolgen. Ferner können ausführlichere Hilfeanfragen per Sprache („Help me find a restaurant“) gestartet werden. Falls die Hilfedomäne nicht spezifiziert wurde, wird ein Dialog initiiert, um die Hilfedomäne zu ermitteln (siehe Dialog 3.2). Wird die Hilfedomäne spezifiziert oder geht diese aus

User:	Help.
System:	I can help you find a restaurant, get subway directions or move around the map. Which one would you like to know about?

Dialog 3.2: MATCH Hilfe

der ursprünglichen Hilfeanfrage hervor (z.B. „Help me find a restaurant“), werden die relevanten Kontextinformationen extrahiert. Das Hilfesystem wertet dazu das momentane Zoom Level (hoch oder niedrig) sowie die Anzahl der auf dem Bildschirm sichtbaren Entitäten (z.B. Restaurants) aus. Ein hohes Zoom Level impliziert beispielsweise, dass viele Restaurants und keine zugehörigen Details auf der Karte zu sehen sind. Eine entsprechende akustische Hilfeausgabe unterstützt den Benutzer dabei, detailliertere Informationen über Restaurants zu erhalten (siehe Dialog 3.3). Befindet sich der Benutzer bereits auf einem niedrigeren Zoom Level,

User:	Show me italian restaurants in Manhattan.
User:	Help.
System:	You can get a closer look at restaurants by circling the area or the restaurant you're interested in and saying, for example, <i>zoom in here</i> , or just writing <i>zoom</i> .

Dialog 3.3: MATCH Hilfe für ein hohes Zoom Level

wird Hilfe zur Abfrage von Detailinformationen gegeben (siehe Dialog 3.4) Neben der Erläuterung verschiedener Interaktionsmöglichkeiten demonstriert das Hilfesystem synchron zur Sprachausgabe auch haptische Eingabemöglichkeiten. Während der Ausgabe der Phrase „zoom in here“ aus Dialog 3.3 zeichnet das Hilfesystem einen Kreis auf einer Kartenregion; während der Anweisung „or just wri-

User: Zoom in here.
 User: Help.
 System: You can get information about restaurants such as the phone number, address or review. For example, to get the phone number, circle a restaurant, like this, and say or write *phone*.

Dialog 3.4: MATCH Hilfe für ein niedriges Zoom Level

te zoom“ schreibt das Hilfesystem das Wort „zoom“ auf dem Display (siehe Abbildung 3.4(a)). Gleiches gilt für Dialog 3.4 (siehe Abbildung 3.4(b)) Die expliziti-

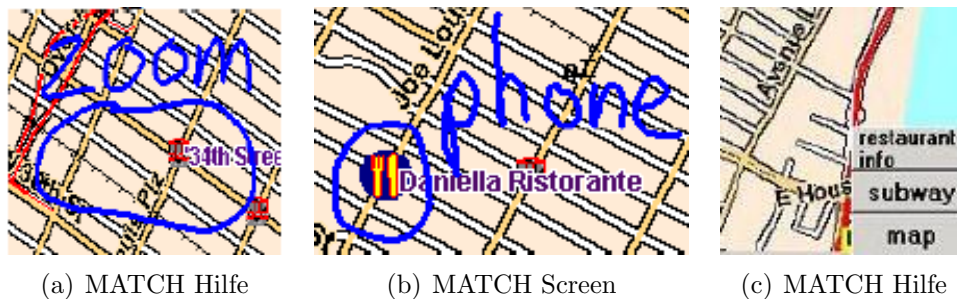


Abbildung 3.4: MATCH System. Quelle: Hastie et. al. (2002)

te Anleitung zu verschiedenen Eingabemöglichkeiten soll den Benutzer dazu bringen, verschiedene Modalitäten zu verwenden und die kognitive Belastung auf diese Modalitäten zu verteilen (Mayer und Moreno, 2003).

In manchen Kontexten werden neben der Hilfestellung zu sprachlichen oder haptischen Eingaben auch Hilfebuttons (Restaurant Info, Subway, Map) eingeblendet, die mit dem Eingabestift angewählt werden können (siehe Abbildung 3.4(c)). Diese Buttons sind nur für einen Dialogschritt sichtbar und bieten dem Benutzer eine fehlerrobuste Eingabevariante.

Die Notation der Hilfedialoge erfolgt über ein XML-Schema, in dem die Prompts und Aktionen für die Hilfeausgaben definiert werden. Die Synchronisation der Aktionen mit den Prompts wird durch Setzen von Bookmarks an den gewünschten Positionen der Sprachausgaben erreicht:

```
<mhdm>
  <speech>
    You can get information about restaurants such as the phone
    number, address or review. For example, to get the phone
    number, circle a restaurant <bookmark>1</bookmark> like
    this, and say or write <bookmark>2</bookmark> phone.
  </speech>
```



```
<actions>
  <action><id>1</id>
    <spec>select(restaurant)</spec>
  </action>
  <action><id>2</id>
    <spec>writeink(phone)</spec>
  </action>
</actions>
</mhdm>
```

3.2.8 Hilfesysteme für SDS in automotiven Umgebungen

Hilfesysteme für SDS in automotiven Umgebungen legen im Vergleich zu solchen für Softwaresysteme eine gesonderte Betrachtungsweise nahe, zudem müssen bei der Entwicklung andere Rahmenbedingungen berücksichtigt werden. Die Primäraufgabe der Fahrzeugführung darf durch FIS nicht beeinträchtigt werden. Dargebotene Informationen müssen daher in verstärktem Maße auf ein Ziel fokussiert sein, um die Ablenkung des Fahrers möglichst gering zu halten. Weiterführende Informationen, z.B. zu ähnlichen Funktionen oder Themengebieten, können sich im Fahrzeug negativ auf die Primäraufgabe auswirken.

Zudem sind die Interaktionsmöglichkeiten des Fahrers mit FIS eingeschränkt. Die Anzeige-Bedienkonzepte der einzelnen Fahrzeughersteller sind proprietär, d.h. Interaktionsparadigmen wie die direkte Manipulation von Objekten auf einer metaphernbasierten graphischen Benutzeroberfläche (GUI; z.B. Microsoft Windows) können nicht auf die Bedienung der FIS übertragen werden. GUIs in Fahrzeugen folgen einer eigenen Bedienlogik, wobei die einzelnen Hersteller unterschiedliche Konzepte verfolgen (siehe Abbildung 3.5(a) und 3.5(b)). Ebenso verschieden wie der Aufbau der GUIs sind die Elemente für die Bedienung derselben. Für die Navigation innerhalb der Menüs stehen dem Benutzer zum einen haptische Bedienelemente (Dreh- und Drücksteller) und zum anderen SDS zur Verfügung (siehe Abbildung 3.5(c) und 3.5(d)). Im Gegensatz zu klassischen Anwendungen aus dem Bereich der Betriebssysteme und Anwendungssoftware, haben sich SDS im Automobil inzwischen etabliert und decken ein breites Spektrum an Funktionen ab. Da die Interaktion zwischen Mensch und Maschine mittels Sprache jedoch immer noch einer Phase der Gewöhnung und Vertrauensbildung bedarf, gelten für die Entwicklung sprachlicher Hilfesysteme auch aus diesem Grund besondere Anforderungen. Erfahrungen aus der Untersuchung zu traditionellen aktiven oder passiven Hilfesystemen können nur bedingt auf die Domäne sprachlicher Hilfesysteme, insbesondere im automotiven Bereich, übertragen werden.



(a) Audi MMI Display. Quelle: Audi (2007)



(b) BMW iDrive Display. Quelle: BMW Group (2007)



(c) Audi MMI Bedienelemente. Quelle: Audi (2007)



(d) BMW iDrive Controller. Quelle: BMW Group (2007)

Abbildung 3.5: GUIs und Bedienelemente verschiedener Anzeige-Bedienkonzepte

Bei der Generierung von Hilfeausgaben verfolgen die Fahrzeughersteller verschiedene Ansätze¹. Das SDS von Audi verfügt zum einen über eine allgemeine Hilfe, welche Hinweise zur Bedienung und für den Aufruf weiterer Hilfethemen enthält (siehe Dialog 3.5). Zudem kann jederzeit Hilfe zu einem Gerät aufgerufen werden.

User:	Hilfe MMI.
System:	Ein kurzer Druck auf die PTT-Taste am Lenkrad aktiviert die Sprachbedienung oder selektiert einen Eintrag bei einer Listenauswahl. Ein langer Druck auf PTT bricht einen laufenden Dialog ab. Sie können bei jeder Eingabe Abbruch oder Hilfe sagen. Für Informationen zu einem Gerät sagen Sie beispielsweise Hilfe Radio oder Hilfe CD usw. Mögliche Geräte sind Radio, CD, Adressbuch, Telefon, Navigation.

Dialog 3.5: AUDI MMI SDS: allgemeine Hilfe

Die Ausgaben sind mehrstufig aufgebaut und enthalten jeweils eine begrenzte Anzahl verfügbarer Kommandos. Bei Bedarf können zusätzliche Informationen angefordert werden (siehe Dialog 3.6). Die kontextspezifische Hilfefunktion bietet dem

¹Die Untersuchung der hier vorgestellten Hilfesysteme fand in den jeweiligen Fahrzeugen statt. Eine zusätzliche Informationsquelle stellen die Bedienungsanleitungen der betrachteten Fahrzeuge dar.

User:	Hilfe Navigation.
System:	Wichtige Befehle für die Navigation sind: Länderauswahl, Zielort, Postleitzahl oder Straße eingeben sowie Navigiere zu einem Adressbucheintrag, für den Sie einen Namen vergeben haben und Zielführung starten bzw. stoppen. Wollen Sie weitere Hilfe?
User:	Ja.
System:	Weitere Befehle für die Navigation sind: Sonderziel, Letzte Ziele, Top Ziele, Info, Dynamisierung ein/aus, Verkehrsinfo und Abbruch. Wollen Sie Hilfe zur Kartendarstellung?
User:	Ja.
System:	Mögliche Kartendarstellungen sind: Zoomautomatik, Übersichtskarte, Detailkarte, Zielkarte, Positionskarte, Maßstab 50m bis 500km sowie Karte größer und Karte kleiner. Allgemeine Hilfe erhalten Sie mit dem Kommando Hilfe MMI.

Dialog 3.6: AUDI MMI SDS: Hilfedialog zu Navigation

Benutzer Informationen zu Kommandos, welche im aktuellen Kontext gesprochen werden können (siehe Dialog 3.7).

User:	Hilfe.
System:	In der Liste blättern können Sie mit weiterblättern, zurückblättern, Listenanfang. Mit dem Kommando Eingabe wiederholen können Sie das Ziel noch einmal eingeben. Mit Postleitzahl können Sie eine Postleitzahl für einen Ort eingeben.

Dialog 3.7: AUDI MMI SDS: Hilfedialog im Kontext Zielliste/Navigation

BMW bietet zwei Hilfefunktionen an, zum einen generelle Hinweise zur Bedienung des Systems (Kommando „Hilfe“), zum anderen dient das Kommando „Optionen“ dazu, eine Liste aller sprechbaren Kommandos des aktuellen Kontexts auszugeben (siehe Dialog 3.8).

User:	Optionen.
System:	Nennen Sie einen Namen gefolgt von anrufen, z.B. Markus Abel anrufen, oder sagen Sie Name wählen, Nummer wählen, Nummer speichern, sowie Telefonbuch, Kurzwahlliste, Top acht, Wahlwiederholung, angenommene Anrufe, entgangene Anrufe, aktive Anrufe und Telefon aus.

Dialog 3.8: iDrive SDS: Hilfedialog im Kontext Telefon

In Hof (2004) und Hof und Hagen (2006) wird ein überarbeitetes Konzept für das iDrive Hilfesystem vorgestellt. Ähnlich wie beim Audi MMI, erfolgt im ent-

wickelten Prototyp eine Strukturierung der Hilfeausgaben hinsichtlich funktionaler Zusammengehörigkeit oder Nutzungshäufigkeit einzelner Funktionen. Beide Optimierungsansätze leisten einen positiven Beitrag zum Umgang mit dem iDrive SDS. Im Zuge einer steigenden Funktionsvielfalt stößt jedoch auch dieses Konzept an seine Grenzen, da die Komplexität der Dialoge innerhalb des Hilfesystems proportional zu der des SDS steigt (siehe Dialog 3.9).

User:	Wahlwiederholung.
System:	Wahlwiederholung. Nennen Sie einen Eintrag oder wählen Sie Liste vorlesen, Liste anzeigen oder Liste löschen.
User:	Optionen.
System:	Optionen. Wählen Sie Listennamen oder Listenfunktionen. Sagen Sie Telefonmenü um ins Menü zurückzukehren.
User:	Listenfunktionen.
System:	Listenfunktionen. Wählen Sie Verwaltungsfunktionen oder Navigationsfunktionen. Sagen Sie Telefonmenü um ins Menü zurückzukehren.
User:	Verwaltungsfunktionen.
System:	Verwaltungsfunktionen. Sagen Sie innerhalb einer Liste anzeigen, vorlesen oder löschen. Um einen Eintrag zu wählen sagen Sie einen Namen oder eine Nummer. Sagen Sie Telefonmenü um ins Menü zurückzukehren.
User:	Liste vorlesen.

Dialog 3.9: iDrive SDS mit überarbeitetem Hilfeconcept: Hilfedialog im Kontext Wahlwiederholung

Ford und Honda verfolgen mit ihren SDS Hilfesystemen einen kontextspezifischen Ansatz. Das Kommando „Hilfe“ listet alle sprechbaren Kommandos im aktuellen grafischen Kontext auf. Honda bietet daneben mit dem Kommando „Voice Command Help“ eine Auflistung aller sprachbedienbaren Funktionen (ca. 200) an.

Das SDS des Hersteller Infiniti weist kein Hilfesystem mit Dialogmöglichkeit auf. Das Kommando „Comand List“ lässt zunächst eine weitere Spezifizierung der Hilfeanfrage zu (z.B. „Phone Help“ oder „Navigation Help“). Nach der Auswahl eines dieser Bereiche wird der Sprachdialog geschlossen, die möglichen Kommandos im aktuellen Kontext werden auf dem Display angezeigt, jedoch nicht vorgelesen.

Lexus stellt lediglich eine statische Hilfe bereit, die ebenfalls nicht sprachbedienbar ist. Das Kommando „Hilfe“ öffnet eine hierarchisch aufgebaute Übersicht möglicher Hilfethemen auf dem Display. Innerhalb der Hilfe kann lediglich mit dem Controller navigiert werden, die Kommandos werden zudem nicht vorgelesen. Unabhängig vom Kontext, in dem die Hilfe aufgerufen wird, öffnet sich die Hilfe stets auf der Übersichtsseite. Die Hilfe gibt Informationen zu kontextunabhängigen und kontextabhängigen Sprachkommandos, jedoch ist bei letzteren z.T.

nicht klar gekennzeichnet, in welchen Kontexten diese gültig sind. Durch Drücken der PTT-Taste wird das Hilfefenster wieder geschlossen.

Mercedes verfügt über eine dreigeteilte Helfefunktionalität. Mit „Hilfe Linguatronic“ können ausführliche Hinweise zum Umgang mit dem SDS aufgerufen werden. „Hilfe Funktionen“ listet die möglichen Geräte wie z.B. Navigation oder Radio auf, „Hilfe“ gibt die sprechbaren Kommandos im aktuellen Kontext aus. Als einziges System ermöglicht Linguatronic eine haptische Interaktion während der Hilfeausgaben. Drückt der Benutzer während der Ausgabe die PTT-Taste, wird die Funktion ausgeführt, die gerade vorgelesen wird.

Mit Ausnahme des Linguatronic Hilfesystems verfügt keines der genannten Systeme im Automobilbereich über multimodale Interaktionsmöglichkeiten. Basierend auf dem in Kapitel 2.1.2 definierten Klassifikationsschema, können die bestehenden Systeme im Sinne einer multimodalen Ausrichtung als exklusive Systeme definiert werden, da die Verwendung der einzelnen Modalitäten nicht oder nur minimal parallel möglich ist und die Interaktionen in verschiedenen Modalitäten nicht miteinander verknüpft sind (siehe Abbildung 3.6).

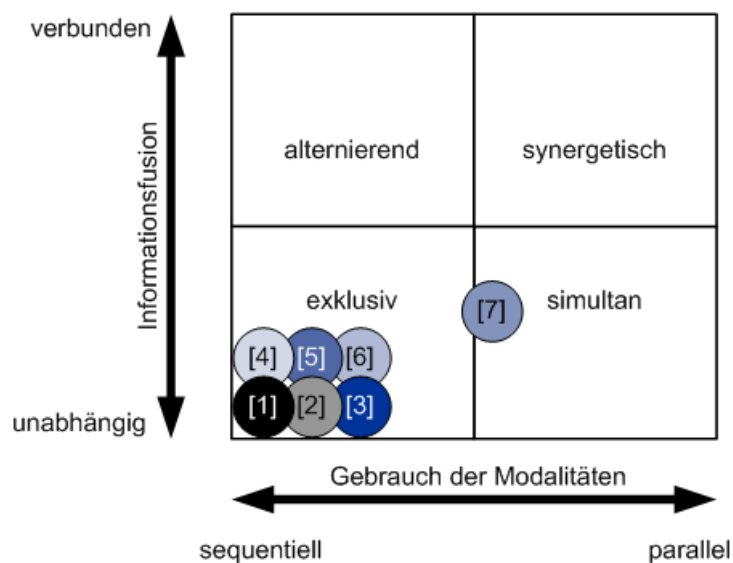


Abbildung 3.6: Klassifikation von Hilfesystemen. [1]: Audi, [2]: BMW, [3]: Ford, [4]: Honda, [5]: Infiniti, [6]: Lexus, [7]: Mercedes

3.3 Fazit

Die vorgestellten Hilfesysteme verfolgen verschiedenste Ansätze zur Generierung von Hilfeausgaben. Abbildung 3.7 gibt nochmals einen Gesamtüberblick über die

Eigenschaften der einzelnen Hilfesysteme. Die Bewertung der einzelnen Hilfesysteme basiert auf den in Kapitel 3.1 erläuterten Merkmalen zur Kategorisierung von Hilfesystemen sowie den in Kapitel 2.1.3 aufgeführten Klassifikationsschemata nach Biemans et. al. (2002) und Kass und Finin (1988).

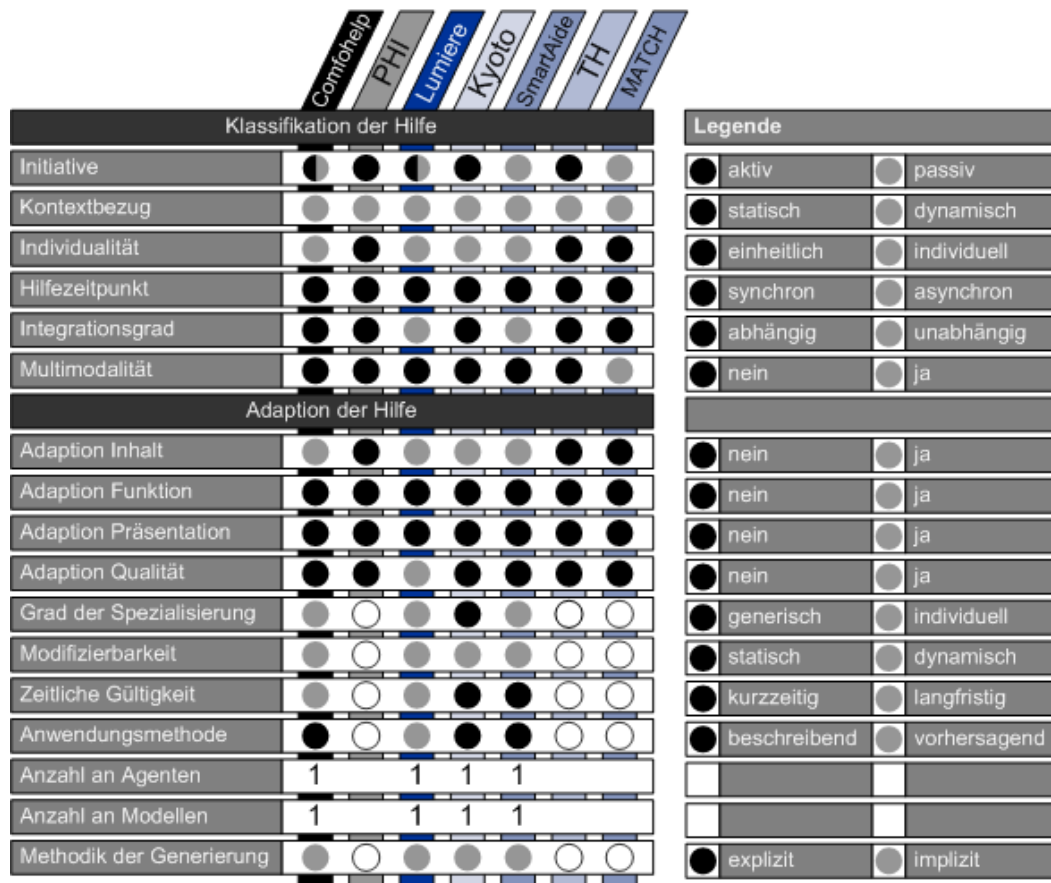


Abbildung 3.7: Klassifikation interaktiver Hilfesysteme

Die in Kapitel 3.2.8 beschriebenen Hilfesysteme für SDS in Fahrzeugen können hinsichtlich ihrer zugrunde liegenden Konzepte differenziert werden (siehe Abbildung 3.8). Alle bestehenden Systeme sind passive Hilfesysteme, d.h. die Hilfe wird explizit durch den Benutzer aufgerufen. Die meisten Hilfesysteme bieten dabei eine kontextspezifische Hilfe an, d.h. ein Benutzer erhält mögliche Sprachkommandos für den Kontext des aktuellen Systemzustands. Einzig die Hilfesysteme der Hersteller Infiniti und Lexus stellen lediglich eine statische Helfefunktionalität zur Verfügung. Alle Hilfesysteme beinhalten eine einheitliche Helfefunktionalität, es erfolgt keine Anpassung der Hilfe an Kenntnisse des Benutzers oder andere Einflussfaktoren. Alle Hilfesysteme sind proprietär und können nur mit den zugehörigen SDS betrieben werden.

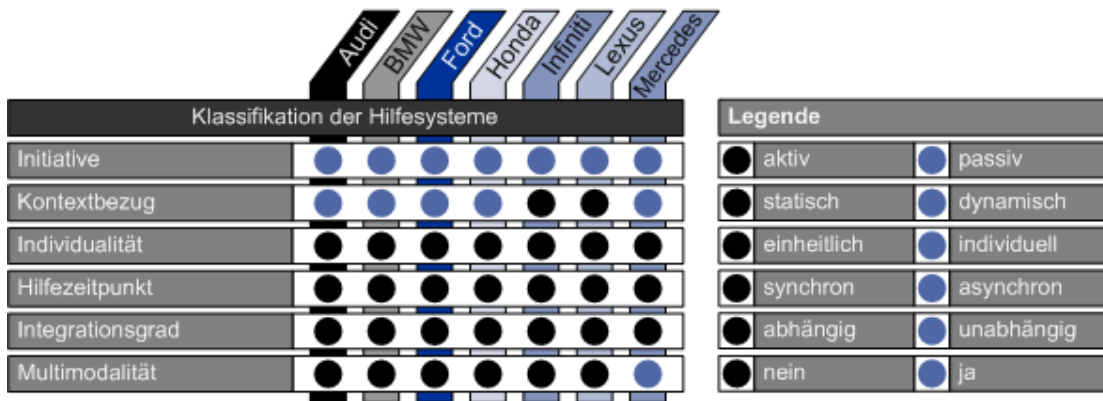


Abbildung 3.8: Klassifikation bestehender Hilfesysteme für SDS in Fahrzeugen

Die aufgeführten Hilfesysteme ermöglichen multimodale Interaktionen, jedoch sind diese mit Ausnahme der Linguatronic von Mercedes weder parallel verwendbar noch miteinander verbunden und können somit nur als exklusiv-multimodale Systeme bezeichnet werden (siehe Abbildung 3.6). Die Linguatronic hingegen ermöglicht während der Sprachausgaben eine haptische Bedienung. Da jedoch keine umfangreicheren Interaktionen möglich sind, kann dieses Hilfesystem nicht als vollwertiges simultan-multimodales System klassifiziert werden.

Kapitel 4

Analyse und Bewertung interaktiver Hilfesysteme

In Kapitel 4.1 werden einleitend allgemeine Anforderungen an ein Hilfesystem für SDS im Fahrzeug erläutert, z.B. wie die Inhalte der Hilfe aufgebaut sein sollen, welche Form der Adaption sinnvoll ist und wie multimodale Interaktionen integriert werden sollen. Neben konzeptuellen Kriterien werden zudem kurz die speziellen Rahmenbedingungen für den Einsatz im Fahrzeug erläutert. Anschließend erfolgt eine Bewertung bestehender Hilfesysteme in Fahrzeugen (Kapitel 4.2) sowie die Beurteilung alternativer Hilfekonzepte aus anderen Anwendungsbereichen (Kapitel 4.3).

4.1 Generelle Anforderungen an Hilfesysteme künftiger Fahrzeuggenerationen

Basierend auf den betrachteten Klassifikationsschemata, muss ein Hilfesystem für SDS in künftigen Fahrzeuggenerationen die Schwachstellen bisheriger Systeme beseitigen. Vor allem unter dem Aspekt einer steigenden Funktionsvielfalt besteht die Herausforderung darin, Hilfesysteme so informativ wie möglich zu gestalten und gleichzeitig die mentale Belastung des Benutzers während der Fahrzeugführung zu minimieren.

Einen wichtigen Beitrag dazu liefern die unter Kapitel 2.3 beschriebenen Kriterien für benutzerfreundliche Sprachdialogsysteme. Ein optimales Hilfesystem muss alle Anforderungen des Kriterienkatalogs erfüllen. Besonderes Augenmerk sollte dabei dem Aspekt der Individualisierbarkeit bzw. Adaption zukommen. Keines der bisher entwickelten Hilfesysteme im automotiven Bereich bietet Mechanismen, mit deren Hilfe ein Benutzer das System anpassen kann bzw. auf deren Basis eine systemseitige Adaption möglich ist. Nach Libuda und Kraiss (2003) stellen insbesondere die Adaption des Systems an Wissen und Präferenzen des Benutzers

sowie die Anpassung an wechselnde Rahmenbedingungen während eines Dialogs vielversprechende Ansätze dar.

Bei der Generierung von Hilfedialogen sollte v.a. der individuelle Anspruch an relevante Informationen berücksichtigt werden. Nach Fischer (2001) können die Kenntnisse eines Benutzers über ein Mensch-Maschine-System in drei Kategorien eingeteilt werden (siehe Abbildung 4.1). Die Menge F1 bildet die Funktionen ab,

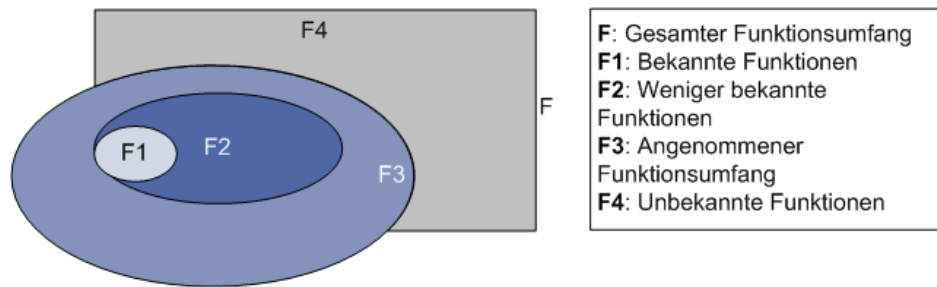


Abbildung 4.1: Kenntnisse eines wenig erfahrenen Benutzers über Systemfunktionen

die dem Benutzer gut bekannt sind, die er häufig nutzt und die einfach zu verwenden sind. F2 repräsentiert Funktionen, die der Benutzer nur vage kennt, selten benutzt und für deren Verwendung er häufiger die Hilfe aufrufen muss. F3 bildet die Menge aller Funktionen, von denen der Benutzer glaubt, dass diese im System enthalten sind. F4 beinhaltet alle Funktionen, die der Benutzer nicht kennt und F bildet die Obermenge aller tatsächlich verfügbaren Systemfunktionen.

Basierend auf diesem Modell sollten die Hilfeinhalte zunächst darauf ausgerichtet sein, den Erstkontakt eines Benutzers mit dem Dialogsystem zu erleichtern und die Menge F1 zu etablieren. Sobald ein Benutzer über genügend Erfahrungen mit einfachen Funktionen gesammelt hat und diese häufig verwendet, kann die Ausrichtung der Hilfe allmählich auf die Funktionen F2 übergehen. Mit zunehmender Verwendung einzelner Funktionen aus F2 können diese F1 zugerechnet werden. Ein weiterer Bestandteil eines Hilfesystems sollte darin bestehen, mittel- bis langfristig das Wissen des Benutzers über die Möglichkeiten des Systems (F3) zu erweitern und F3 möglichst mit F4 in Einklang zu bringen. Auf diese Weise etabliert sich mit steigendem Erfahrungsgrad eine stabile Menge F1 und F2 sowie im Optimalfall Wissen über die Möglichkeiten und Grenzen des Systems (siehe Abbildung 4.2).

Hinsichtlich der eingesetzten Interaktionsparadigmen sollte der multimodalen Interaktion eine erhöhte Bedeutung zukommen. Mit Ausnahme der Linguatronic von Mercedes sind alle automotiven Hilfesysteme exklusiv-multimodale Systeme. Da im Fahrzeug die primäre Aufgabe stets die Fahrzeugführung ist, müssen die sekundären Aufgaben (Bedienung der FIS und FAS) bestmöglich an die Primäraufgabe angepasst werden. Nach Neuss (2001) und Niedermaier (2003) bieten

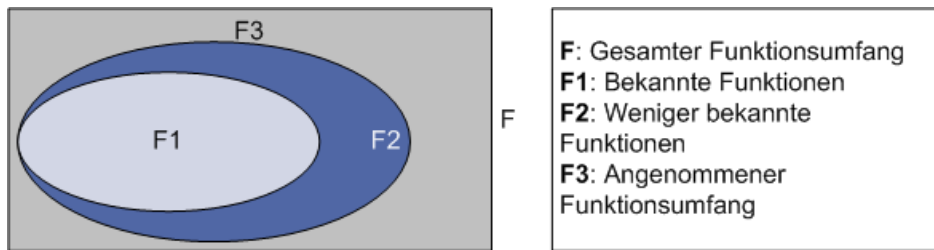


Abbildung 4.2: Kenntnisse eines erfahrenen Benutzers über Systemfunktionen

alternierend-multimodale Dialogsysteme in Kraftfahrzeugen das größte Potenzial, Benutzer optimal bei der Erreichung eines Ziels zu unterstützen und gleichzeitig die mentale Belastung gering zu halten. Bei synergetisch-multimodalen Systemen kann der Koordinationsaufwand, insbesondere bei der Verwendung verschiedener Eingabemodalitäten, die mentale Belastung des Benutzers unnötig erhöhen. Zudem besitzen nach Oviatt et. al. (1997) synergetisch-multimodale Systeme ihre Stärken hauptsächlich in räumlich zuordnenden Aufgabentypen, wie z.B. der Markierung eines Kartenausschnitts per Touchscreen und paralleler Spracheingabe. Da Aufgaben dieser Art im Fahrzeug nicht auftreten, ist die Entwicklung synergetisch-multimodaler Dialogsysteme in dieser Domäne wenig sinnvoll. Multimodale Systeme sollen demnach je nach Aufgabentyp die optimale Modalität zur Erreichung eines Ziels zur Verfügung stellen (exklusiv-multimodal) oder einzelne Modalitäten sinnvoll miteinander verknüpfen (alternierend-multimodal bzw. simultan-multimodal; siehe Abbildung 4.3).

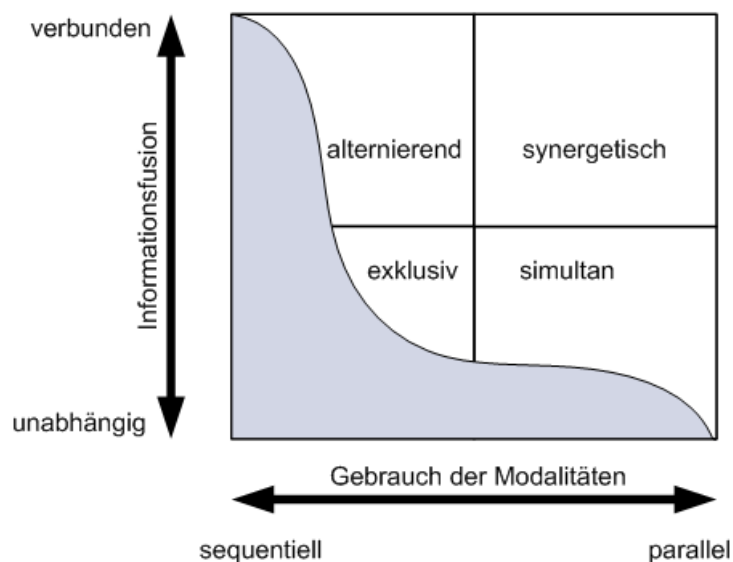


Abbildung 4.3: Anforderung an multimodale Dialogsysteme im Kraftfahrzeug

Neben konzeptuellen Anforderungen muss ein Hilfesystem den speziellen Rahmenbedingungen im Fahrzeug Rechnung tragen. Da die Rechenleistung einer Headunit¹ begrenzt ist und mehrere Applikationen darauf parallel betrieben werden (z.B. Navigationssystem, Entertainment), muss das SDS Hilfesystem extrem performant sein, um einen Dialog in Echtzeit gewährleisten zu können. Ferner muss der Speicherbedarf des Hilfesystems möglichst gering sein, da für die Integration zusätzlicher Systeme nur begrenzte Ressourcen zur Verfügung stehen. Zudem muss das Hilfesystem einfach zu implementieren und zu erweitern sein, da mit jeder Modellüberarbeitung und über verschiedenen Baureihen hinweg unterschiedliche Spezifikationen und Funktionsumfänge für das SDS definiert werden. Als Richtwert für den maximalen Funktionsumfang des SDS müssen ca. 850 Funktionen berücksichtigt werden.

4.2 Bewertung bestehender Hilfesysteme in Fahrzeugen

Die Bewertung der Hilfesysteme kann anhand der in Kapitel 2.3 vorgestellten Richtlinien für den Entwurf von Dialogsystemen erfolgen. Die DIN EN 9241-10, die Acht Goldenen Regeln des Interface Designs sowie die Richtlinien für die Gestaltung von Sprachdialogen überschneiden sich z.T. in ihrer Bedeutung. Aus diesem Grund ist es möglich, diese Richtlinien zu kombinieren und die Hilfesysteme anhand des resultierenden Kriterienkatalogs zu bewerten:

- Selbstbeschreibungsfähigkeit, Feedback (EN ISO 9241-10, Acht Goldene Regeln des Interface Designs): Das System ist intuitiv verständlich, die Rückmeldungen für jeden Dialogschritt unterstützen den Benutzer bei der Verwendung des Systems.
- Aufgabenangemessenheit (EN ISO 9241-10): Das Hilfesystem unterstützt den Benutzer bei der Erledigung seiner Aufgabe.
- Lernförderlichkeit (EN ISO 9241-10): Durch die Verwendung der Hilfe entsteht beim Benutzer eine Vorstellung vom Aufbau und den Möglichkeiten des SDS.
- Steuerbarkeit, Kontrolle (EN ISO 9241-10, Acht Goldene Regeln des Interface Designs): Der Benutzer hat die Möglichkeit, einen Dialog zu starten und innerhalb des Hilfedialogs zu navigieren, bis sein Ziel erreicht ist.
- Individualisierbarkeit, Nutzerwissen (EN ISO 9241-10, Richtlinien für die Gestaltung von Sprachdialogen): Das System stellt Mechanismen zur

¹Die Head Unit bündelt im Fahrzeug alle Fahrerinformationssysteme und Geräte, die nicht zur Fahrzeugführung beitragen (z.B. Telefon, CD Player, Navigationssystem).

benutzer- oder systemseitigen Anpassung an Wissen und Präferenzen des Benutzers zur Verfügung.

- Erwartungskonformität, Konsistenz, Transparenz (EN ISO 9241-10, Acht Goldenen Regeln des Interface Designs, Richtlinien für die Gestaltung von Sprachdialogen): Innerhalb des Systems werden konsistente Dialoge, Farben, Begriffe verwendet.
- Fehlertoleranz, Fehlerbehandlung (EN ISO 9241-10, Acht Goldene Regeln des Interface Designs): Treten Fehler innerhalb eines Dialogs auf, werden diese abgefangen und der Benutzer wird bei der Fortführung des Dialogs unterstützt.
- Geringe KZG-Belastung, kurze Prompts (Acht Goldene Regeln des Interface Designs, Richtlinien für die Gestaltung von Sprachdialogen): Der Informationsgehalt der Systemausgaben werden möglichst gering gehalten.
- Reversibel (Acht Goldene Regeln des Interface Designs): Aktionen sind leicht rückgängig zu machen.
- Shortcuts (Acht Goldene Regeln des Interface Designs): Das Hilfesystem bietet Shortcuts für häufig verwendete Funktionen an.
- Geschlossenheit (Acht Goldene Regeln des Interface Designs): Sprachausgaben bestehen aus einem Anfang (z.B. einer einleitenden Phrase), einem Mittelteil und einem Ende.
- Redundanz (Richtlinien für die Gestaltung von Sprachdialogen): Es treten keine unnötigen oder überflüssigen Interaktionen auf.

Die Bewertung der Hilfesysteme erfolgte anhand des aufgestellten Kriterienkatalogs durch Expertenreviews. Die Vorgehensweise wird im Folgenden exemplarisch durch die Bewertung der Hilfesysteme von BMW und Lexus verdeutlicht, da die Unterschiede hier sehr deutlich werden.

- Selbstbeschreibungsfähigkeit, Feedback: Die Hilfedialoge beim BMW SDS beginnen stets mit einer Rückmeldung der aktuellen Position innerhalb des Dialogs („Optionen. Sagen Sie ...“). Anschließend werden die sprechbaren Kommandos im aktuellen Kontext ausgegeben (Konzept „speak what you see“). Kriterium: erfüllt.

Das Lexus SDS beendet nach Aufruf der Hilfe den laufenden Sprachdialog. Stattdessen wird ein grafisches Hilfemenü auf dem Display angezeigt, das stets auf der obersten Hierarchieebene beginnt. Der Benutzer erkennt zwar, dass die Hilfe aktiviert ist, die sprechbaren Kommandos im Kontext müssen jedoch explizit gesucht werden. Kriterium: teilweise erfüllt.

- Aufgabenangemessenheit: Das iDrive SDS listet nach dem Hilfeaufruf die kontextspezifischen Sprachkommandos auf und unterstützt den Benutzer somit beim Erreichen eines Ziels. Kriterium: erfüllt.

Da Lexus keine kontextspezifische Hilfe anbietet, müssen die sprechbaren Kommandos erst in der Hilfe gesucht werden, was eine stärkere mentale Belastung und eine erhöhte Ablenkung des Fahrers zur Folge haben kann. Die sprechbaren Kommandos sind zwar hinterlegt, die Navigation innerhalb der Hilfe ist jedoch suboptimal. Kriterium: teilweise erfüllt.

- Lernförderlichkeit: Die Verwendung der iDrive-Hilfe leistet einen wichtigen Beitrag für die Erlernbarkeit von Sprachkommandos. Untersuchungen zum Lernverhalten zeigen, dass die Ausgabe akustischer Informationen eine gute Hilfestellung für den Lernerfolg der Benutzer bieten (vgl. Kapitel 5.4). Kriterium: erfüllt.

Im Gegensatz dazu gestaltet sich die Suche nach einem bestimmten Kommando beim Lexus weniger intuitiv. Zwar stellt sich auch hier ein Lerneffekt ein, jedoch ist dieser mit einem größerem Aufwand verbunden. Kriterium: teilweise erfüllt.

- Steuerbarkeit, Kontrolle: Das BMW System erlaubt es nicht, innerhalb der Hilfe zu navigieren. Unabhängig von der Länge der Dialoge, werden alle Sprachkommandos konsekutiv ausgegeben. Die Ausgabe kann lediglich wiederholt werden, jedoch ist dies nur mit dem Kommando „Optionen“ möglich, ein Kommando „wiederholen“ existiert nicht. Durch den Aufruf eines Sprachkommandos aus den Optionen kann in den regulären Systemdialog gewechselt werden. Kriterium: teilweise erfüllt.

Das Lexus SDS bietet keine sprachlichen Hilfedialoge an. Sobald die Hilfe aufgerufen wurde, schließt sich der Sprachdialog. Nach Verwendung der Hilfe muss der ursprüngliche Dialog erneut gestartet werden. Kriterium: nicht erfüllt.

- Individualisierbarkeit, Nutzerwissen: Keines der beiden Hilfesysteme bietet Mechanismen zur Anpassung der Dialoge an Wünsche, Präferenzen oder Kenntnisse des Benutzers. Kriterium: nicht erfüllt.
- Erwartungskonformität, Konsistenz, Transparenz: Jeder Hilfedialog beginnt mit der Phrase „Optionen“ und die einzelnen Hilfedialoge des iDrive-Systems beinhalten generell die Kommandos, die in einem bestimmten graphischen Kontext sprechbar sind („speak what you see“). Die Inhalte der Hilfe sind in den verschiedenen Kontexten stets gleich. Kriterium: erfüllt.

Die Erwartung, dass der Aufruf der Hilfe in einer konkreten Dialogsituation Hilfestellung leistet, wird enttäuscht. Die kontextunabhängige Hilfe ist jedoch zu großen Teilen eindeutig und die jeweiligen Kommandos enthalten durchgängig Informationen darüber, in welchen Kontexten sie gesprochen werden können. Kriterium: teilweise erfüllt.

- Fehlertoleranz, Fehlerbehandlung: Die Fehlerbehandlung wird innerhalb des Hilfesystems beim BMW nur rudimentär behandelt. Auftretende Erkennen-

fehler werden mit der Bitte der nochmaligen Spracheingabe abgefangen. Kriterium: teilweise erfüllt.

Da Lexus keine sprachbedienbare Hilfe besitzt, ist dieses Kriterium hier nicht erfüllt.

- Geringe KZG Belastung, kurze Prompts, geringe mentale Belastung: Das BMW System beinhaltet in manchen Kontexten Hilfeausgaben, die mehr als neun Kommandos beinhalten bzw. länger als 25 Sekunden dauern. Die Fähigkeit, alle Kommandos im Gedächtnis zu behalten und nach der Hilfeausgabe zu verwenden, ist eingeschränkt. Da jedoch nur ein Teil der Hilfedialoge betroffen ist, ist das Kriterium teilweise erfüllt.

Die mentale Belastung der Benutzers bei der Verwendung der Hilfe, insbesondere während der Fahrt, ist sehr hoch. Die Konzentration des Benutzers richtet sich von der Straße auf das Display, für die Navigation innerhalb der Hilfe muss ein großer Teil der Aufmerksamkeit auf die Sekundäraufgabe gerichtet werden. Kriterium: nicht erfüllt.

- Reversibel: Keines der beiden System stellt Mechanismen zur Verfügung, mit deren Hilfe Aktionen innerhalb der Hilfe rückgängig gemacht werden können. Kriterium: nicht erfüllt.
- Shortcuts: Aus der Hilfe des BMW Systems heraus können sämtliche Shortcuts verwendet werden, die auch in den regulären Dialogen als Shortcut angeboten werden. Kriterium: erfüllt.

Mangels Dialog erfüllt Lexus dieses Kriterium nicht.

- Geschlossenheit: Alle Sprachausgaben der Hilfe beginnen mit „Optionen“, wonach eine Liste mit möglichen Sprachkommandos vorgelesen wird. Die Prosodie der Sprachausgaben gibt zudem Aufschluß über das Ende der Sprachausgabe. Kriterium: erfüllt.

Lexus besitzt keine Sprachdialoge, daher ist auch dieses Kriterium nicht erfüllt.

- Redundanz: Die Hilfe des iDrive SDS formuliert Hilfeausgaben speziell für einen Kontext. Die Information zum Erreichen eines Ziels innerhalb des Kontexts wird somit unmittelbar angeboten. Kriterium: erfüllt.

In der Hilfe des Lexus Systems muss zunächst nach den entsprechenden Kommandos gesucht werden. Kriterium: nicht erfüllt.

Die Bewertung der restlichen Hilfesysteme anhand dieser Vorgehensweise zeigt, dass das Hilfesystem des Audi MMI den anderen Systemen überlegen ist (siehe Abbildung 4.4). Besonders die Aufteilung der Hilfe in verschiedene Ebenen und die damit einhergehende Reduktion des Informationsgehalts wirken sich positiv auf das Ergebnis aus.

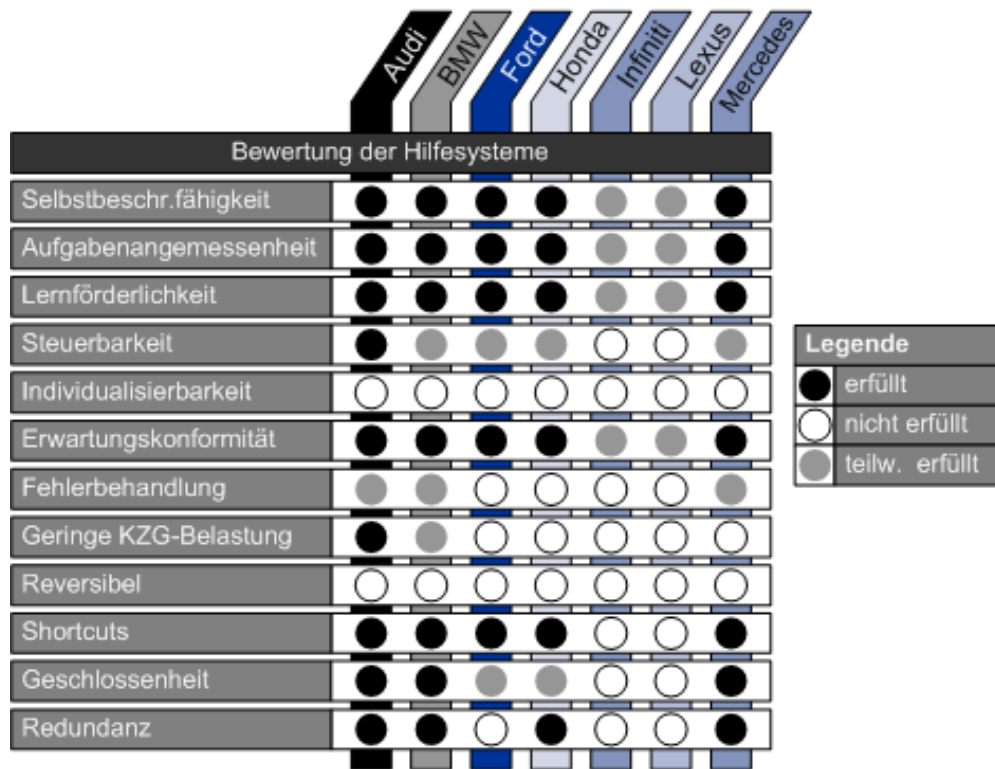


Abbildung 4.4: Bewertung von Hilfesystemen in Fahrzeugen

Die Hilfekontexte von BMW und Mercedes sind nahezu identisch. Durch die fehlende Strukturierung der Hilfeaushaben kommen jedoch z.T. sehr lange Systemausgaben zustande. Besonders negativ fällt dabei das Linguatronic Hilfesystem auf, welches in manchen Kontexten mehr als 20 Systemfunktionen zur Auswahl anbietet. Im Gegensatz zum iDrive Hilfesystem sind jedoch bei der Linguatronic hinsichtlich der Multimodalität simultan-komplementäre Interaktionen möglich.

Die Hilfesysteme von Ford und Honda offerieren ebenfalls kontextspezifische Hilfe, listen die Kommandos jedoch nur auf. Daher kommt kein natürlich anmutender Dialog zustande. Die Auflistung sämtlicher Sprachkommandos des SDS im Honda hat keinen erkennbaren Mehrwert.

Infiniti und Lexus stellen keine kontextspezifischen Hilfesysteme zur Verfügung. Die Aktivierung einer statischen, visuellen Hilfe reduziert den Mehrwert der Sprachbedienung. Besonders während der Fahrt ist diese Art der Hilfe als kontraproduktiv zu bewerten.

4.3 Bewertung alternativer Hilfekonzepte

Im Gegensatz zu bereits verfügbaren Hilfesystemen in Fahrzeugen, muss bei der Bewertung von Hilfekonzepten aus anderen Anwendungsdomänen zunächst geprüft werden, ob ein Konzept die grundlegenden Anforderungen für den Einsatz im Fahrzeug erfüllt.

Der planbasierte Ansatz des Hilfesystems COMFOHELP (vgl. Kapitel 3.2.1) beruht auf dem Abgleich einer beobachteten Aktionssequenz mit einer Planbibliothek. Dieser einfache Mechanismus wäre generell auf ein SDS übertragbar, das Konzept des Hilfesystems ist jedoch mit einem entscheidenden Nachteil behaftet. Die Planbibliothek muss für jedes Dialogziel explizit definiert werden. Ausgehend von ca. 850 Systemfunktionen fiel die Planbibliothek sehr umfangreich aus und wäre in der von Krause et. al. (1993) vorgeschlagenen Notation umständlich zu pflegen. Da die Menüstruktur eines Anzeige-Bedienkonzepts z.T. sehr komplex ist, müssten für einzelne Dialogziele eine Vielzahl an Plänen definiert werden. Die Änderung eines Kriteriums für die Berechnung einer Route (z.B. „Kurze Route“) umfasst beispielsweise im günstigsten Fall zwei Sprachkommandos („Routenkriterien“, „Kurze Route“), im längsten Fall jedoch fünf Kommandos („Navigation“, „Auswahl“, „Karte“, „Routenkriterien“, „Kurze Route“). Für jeden möglichen Einsprungpunkt im Dialog müsste ein separater Plan hinterlegt werden. Zudem müssten die einzelnen Durchlaufkommandos definiert werden. Dazu gehören neben allgemeinen Funktionen wie der Hilfe auch kontextspezifische Funktionen und Funktionen wie z.B. die Bedienung der Entertainmentsysteme oder des Telefons. Da das Konzept dieses Hilfesystems in komplexen Domänen v.a. in puncto Handhabung Schwächen aufweist, ist es für den Einsatz im Fahrzeug nicht geeignet.

Das planbasierte Hilfesystem PHI (vgl. Kapitel 3.2.2) erzeugt die Planbibliothek zur Laufzeit automatisch mit Hilfe einer modalen Temporallogik aus vordefinierten abstrakten Plänen. Somit umgeht dieser Ansatz zwar die Problematik der expliziten Definition einer Planbibliothek, jedoch wurde das Hilfesystem nur exemplarisch in einer sehr kleinen Anwendungsdomäne mit weniger als zehn Funktionen implementiert. Die Definition abstrakter Pläne in diesem Umfang ist mit einem überschaubaren Aufwand verbunden, in einer komplexen Anwendungsdomäne mit ca. 850 Funktionen nähme der Aufwand erheblich zu. Hinsichtlich der Performanz des Systems liegen keine Untersuchungen vor, jedoch besteht die Möglichkeit, dass die Prozesse der Planerkennung und Plangenerierung auf Basis der Temporallogik in einer komplexen Anwendungsdomäne sehr rechenintensiv sind. Ein weiterer Nachteil des Hilfekonzepts ist die Ausrichtung auf eine aktive Hilfe. Die Rahmenbedingungen dieser Arbeit orientieren sich an den Vorgaben der BMW Group. Die internen Richtlinien für den Entwurf von Sprachdialogen beschreiben hinsichtlich der Systeminitiative eine zurückhaltende Strategie. Der Fahrer soll möglichst keine unaufgeforderten systeminitiierten Dialoge erhalten. Aus diesen Gründen scheint eine Erweiterung dieses Hilfekonzepts wenig sinnvoll.

Das auf einem Bayes'schen Netz basierende Lumière Hilfesystem (vgl. Kapitel 3.2.3) findet in einer reduzierten Form bereits Anwendung im Microsoft Office Assistenten. Die Erfahrungen mit dem Assistenten spiegeln eine zentrale Schwäche des System wider. Da die Berechnung der Hilfethemen auf Wahrscheinlichkeitsannahmen beruht, werden die Aktionen des Benutzers hin und wieder falsch interpretiert und führen zu irrelevanten Hilfeausgaben. Vor allem die Qualität der aktiven Hilfe leidet unter diesem Phänomen. Viele Office Benutzer kennen die Situation, dass sich der Hilfeassistent unaufgefordert aktiviert und Hilfe zu Themen anbietet, die u.U. vom eigentlichen Dialogziel abweichen. Gerade in der Domäne der SDS im Fahrzeug ist jedoch die Erwartungskonformität des Hilfesystems ein entscheidender Erfolgsfaktor. Verhält sich ein System bei den ersten Kontakten nicht wie gewünscht, führt dies schnell zu Frustration und einer allgemeinen Ablehnung des Systems. Die Abbildung eines Hilfesystems für SDS auf Wahrscheinlichkeitsnetzen ist im Fahrzeug somit nur bedingt brauchbar.

Das Hilfesystem des Kyoto Sightseeing Guide (vgl. Kapitel 3.2.4) stellt einen interessanten Ansatz dar, da die Hilfe direkt mit den zugrunde liegenden Konzepten des Dialogsystems (Domain Concept Tree) verknüpft ist. Das Hilfesystem berücksichtigt zudem Erfahrungen, die ein Benutzer im Umgang mit dem System sammelt. Allerdings basiert die Modellierung von Lerneffekten und der Ausbildung von Transferwissen lediglich auf theoretisch nicht fundierten Annahmen über das mögliche Lernverhalten. Des weiteren wurde dieses Hilfesystem konstruiert, um dem Benutzer in einer bestimmten Dialogsituation konkrete Hilfestellung anzubieten. Das Hilfesystem gibt lediglich indirekt in Form der Templates einen Überblick über sprechbare Kommandos. Hier offenbart sich jedoch eine weitere Schwäche des Systems. Falls bei der Berechnung der Hilfe Knoten n_x mit gleichen Werten p_{n_x} in Betracht gezogen werden, erfolgt die Auswahl der Hilfe zufällig. In der komplexen Domäne der SDS im Fahrzeug kann dieser Fall jedoch gehäuft auftreten, weshalb die zufällige Generierung einer Hilfeausgabe einen inakzeptablen Umstand darstellt.

Das SmartAidè Hilfekonzent (vgl. Kapitel 3.2.5) basiert ähnlich wie COMFO-HELP auf einer Aktionsbibliothek, die explizit definiert werden muss. Aus diesem Grund ist dieses Konzept aus oben erläuterten Gründen nicht praktikabel für den Einsatz im Fahrzeug. Ferner beruht das Konzept auf der Idee, visuell-haptische Interaktionen zu demonstrieren. Das Hilfesystem zeigt somit nur Abläufe verschiedener Interaktionen, Zusammenhänge oder zugrunde liegende Funktionsprinzipien bleiben dem Benutzer verschlossen.

Der Targeted Help Ansatz (vgl. Kapitel 3.2.6) ist in seiner grundsätzlichen Auslegung als Zusatz zu einem SDS gedacht. Hilfeausgaben erfolgen nicht per Hilfeaufforderung, sondern im Fall eines schlechten Erkennungsergebnisses des primären Spracherkenners. Die Hilfe erläutert nur indirekt mögliche Sprachkommandos, ein Überblick über sprechbare Kommandos oder Kommandostrukturen ist nicht gegeben. Insofern ist das System als Ergänzung zu einem SDS sinnvoll, für den Erstkontakt

eines Benutzers mit dem SDS liefert die Hilfe jedoch keinen Überblick über das System und seine Eigenschaften. Zudem liegen keine Ergebnisse über die Performance des SLM Spracherkenners in komplexen Domänen vor, es ist jedoch anzunehmen, dass diese nicht das Niveau der eingeschränkten Domäne (wie in Hockey et. al. (2003) beschrieben) erreicht.

Das synergetisch-multimodale Hilfesystem MATCH (vgl. Kapitel 3.2.7) legt den Schwerpunkt der Hilfe auf die Kombination verschiedener Modalitäten. Obgleich die Eigenschaften der Hilfe in ihrer Anwendungsdomäne überzeugen, können die Konzepte nicht ohne Anpassung auf die Fahrzeugdomäne transferiert werden. Zum einen fehlt im Fahrzeug eine entscheidene Interaktionsmöglichkeit (Touchscreen), zum anderen sind die Use Cases des Prototyps äußerst simpel. In komplexen Dialogstrukturen können als einzige Schlüsselkonzepte die Aufteilung der Informationen auf verschiedene Modalitäten und die Synchronisation der Modalitäten bestehen.

4.4 Fazit

Hilfesysteme in künftigen Fahrzeuggenerationen müssen sich an die Bedürfnisse einzelner Benutzer individuell anpassen. Dazu müssen automatisch diejenigen Funktionen eines SDS ermittelt werden, für die der Benutzer keine Hilfe mehr benötigt. Der Hilfefokus liegt auf Funktionen, die ein Benutzer selten verwendet oder die er noch nicht kennt. Hinsichtlich der im Fahrzeug zum Einsatz kommenden Interaktionsparadigmen muss der multimodalen Interaktion eine verstärkte Bedeutung zukommen. Dabei ist es wichtig, je nach Dialogziel die optimale Modalität zur Verfügung zu stellen (exklusiv-multimodal) oder verschiedene Modalitäten sinnvoll miteinander zu verknüpfen (alternierend-multimodal bzw. simultan-multimodal).

Die Bewertung vorhandener Hilfesysteme für SDS in Fahrzeugen offenbart hinsichtlich der in Kapitel 2.3 dargestellten Gestaltungsaspekte für den Entwurf von Sprachdialogen zahlreiche Schwächen. Insbesondere die fehlende Individualisierbarkeit von Dialogen und fehlende oder nur rudimentär vorhandene multimodale Komponenten kennzeichnen Verbesserungspotenziale aktueller Hilfesysteme von SDS. Hilfekonzepte aus anderen Anwendungsbereichen können jedoch aus verschiedenen Gründen nicht. Planbasierte Ansätze, wie sie in COMFOHELP oder PHI verwendet werden, verursachen einen hohen Bereitstellungs- und Wartungsaufwand und benötigen zudem relativ viel Systemressourcen. Das Konzept der Targeted Help kann lediglich als Backupsystem verstanden werden, welches im Falle einer Fehlerkennung durch einen grammatikbasierten Spracherkenner auf einen zweiten Spracherkenner mit statistischem Sprachmodell zurückgreift. Der Benutzer gewinnt keinen Überblick über die Systemfunktionen, für eine passive Hilfefunktion ist Targeted Help somit nicht geeignet. Das Hilfesystem für den Kyoto Sightseeing Guide verfolgt einen interessanten Ansatz, indem die Struktur

des Hilfesystems der Struktur des SDS angepasst wird. Die Ausgabe der Hilfetexte ist jedoch aufgrund seiner Templatestruktur sehr statisch, zudem zeigen sich bei der Generierung der Hilfeinhalte konzeptuelle Schwächen. Das multimodale Hilfesystem MATCH basiert auf der Interaktion eines Benutzers mit einem System via Stift, Touchscreen und Sprache. Eine Übertragung der MATCH-Konzepte auf die Fahrzeugdomäne ist somit nicht oder nur schwer möglich. Für Fahrzeuge künftiger Baureihen müssen demnach Hilfekonzepte entwickelt werden, die auf ressourcensparende Art und Weise den Anforderungen komplexer SDS gerecht werden und sowohl adaptive als auch multimodale Komponenten in sich vereinen.

Kapitel 5

Modellierung eines adaptiven, multimodalen Hilfesystems

Die Bewertung bestehender Hilfesysteme in automotiven Umgebungen in Kapitel 4.2 hat diverse Schwachstellen sowohl konzeptueller Art als auch auf technischer Ebene offengelegt. Die Anforderungsanalyse aus Kapitel 4.1 bildet unter Berücksichtigung kognitionspsychologischer Aspekte (vgl. Kapitel 2.3.2) das Fundament für die Entwicklung eines optimierten Hilfesystems für SDS.

Kapitel 5.1 behandelt zunächst die allgemeine Hilfe. Der kontextspezifischen Hilfefunktionalität kommt im Rahmen dieser Arbeit die größte Bedeutung zu (Kapitel 5.2). Die Ausprägung der Adaption hinsichtlich Dialogstrategie und Hilfeinhalte wird in Kapitel 5.2.2 und 5.2.1 erläutert. Die Integration multimodaler Dialogkomponenten ergänzt die sprachliche Hilfe (Kapitel 5.2.3).

Daneben bildet die kontextunabhängige Hilfe die dritte Säule des Hilfesystems (Kapitel 5.3). Benutzer erhalten hier die Möglichkeit, grundlegende Systemfunktionen mittels natürlichsprachlicher Hilfeanfragen kennen zu lernen. Diese Form der Hilfe ist primär für den Erstkontakt bzw. unerfahrene Benutzer gedacht.

Die zur Implementierung notwendigen Parameter bezüglich des Lernens und Vergessens von Sprachkommandos werden in Kapitel 5.3 auf Basis von Experimenten ermittelt. Die Ergebnisse dieser Untersuchungen bestimmen abschließend die Auswahl der Beschreibungsmodelle für das adaptive Hilfesystem (Kapitel 5.5).

5.1 Allgemeine Hilfe

Das Konzept der allgemeinen Hilfe beinhaltet bei BMW Hinweise zur generellen Bedienung des Systems. Da dieses Konzept auch in künftigen Fahrzeuggenerationen beibehalten werden soll, können an dieser Stelle Informationen zur kontextspezifischen Hilfe untergebracht werden. Zudem thematisiert die allgemeine Hilfe das benutzeradaptive Verhalten der kontextspezifischen Hilfe sowie Metakommandos zur Steuerung von Dialogen. Der Hilfedialog fällt relativ umfangreich aus, da

dieser jedoch grundlegende Themen behandelt, die erfahrungsgemäß nach kurzer Zeit selten bis gar nicht mehr benötigt werden, kann ein längerer Dialog in Kauf genommen werden (siehe Dialog 5.1). Während der sprachlichen Ausgabe der allgemeinen Hilfe unterstützt zudem eine visuelle Hilfetafel auf dem Display den Benutzer dabei, einen Überblick über die sprechbaren Kommandos zu erhalten (vgl. Kapitel 5.2.3).

User:	Hilfe.
System:	Hilfe. Um Hilfe zu Sprachkommandos zu erhalten, sagen Sie Optionen. Die Optionen beinhalten Informationen zu sprechbaren Kommandos in verschiedenen Kontexten. Ferner passen sich die Optionen mit der Zeit an Ihr Bedienverhalten an. Sie können die Anpassungen jederzeit im Menü Einstellungen rückgängig machen. Innerhalb der Hilfe stehen Ihnen die Kommandos Weiter, Zurück und Wiederholen zur Verfügung. Falls Sie einen Dialog beenden möchten, sagen Sie Abbruch.

Dialog 5.1: Allgemeiner Hilfedialog

5.2 Kontextspezifische Hilfe

Die kontextabhängige Hilfe bildet das Rückgrad der meisten SDS in automatisierten Umgebungen (vgl. Kapitel 3.2.8). Sie bietet einem Benutzer gezielt Informationen zu sprechbaren Kommandos innerhalb eines Kontexts und unterstützt den Benutzer kurzfristig beim Erreichen eines Dialogsziels. Mittel- und langfristig trägt die kontextspezifische Hilfe auch zum Erlernen der möglichen Sprachkommandos bei. Besonders problematisch sind jedoch häufig die sehr langen Hilfeausgaben, die Linguatronic von Mercedes enthält beispielsweise in manchen Kontexten mehr als 25 Sprachkommandos. Die einzelnen Systemausgaben sollten jedoch nicht mehr als $4(\pm 2)$ Kommandos beinhalten, um den Primacy- und Recency-Effekt bei der kognitiven Verarbeitung optimal auszunutzen (siehe Abbildung 5.1; vgl. Kapitel 2.3.2).

Die Ausgabe der möglichen Sprachkommandos kann folglich nicht mehr in einer einzelnen Sequenz erfolgen, sondern muss auf verschiedene Dialogschritte aufgeteilt werden. Dies trägt zur Verkürzung der Sprachausgaben und zur besseren Memorierungsfähigkeit der Kommandos für einzelne Dialogschritte bei. Die Anzahl der Dialogschritte ist abhängig von der Anzahl der Sprachkommandos pro Dialogschritt. In Salmen (2003) werden Ausgaben bis zu neun Kommandos pro Dialogschritt diskutiert, jedoch zeigt die Untersuchung von Systemausgaben im iDrive Hilfesystem, dass die Benutzer mehr als sechs Kommandos als zu lang empfinden (Hof, 2004). Auf Basis dieser Ergebnisse erscheint die Verwendung von $4(\pm 2)$

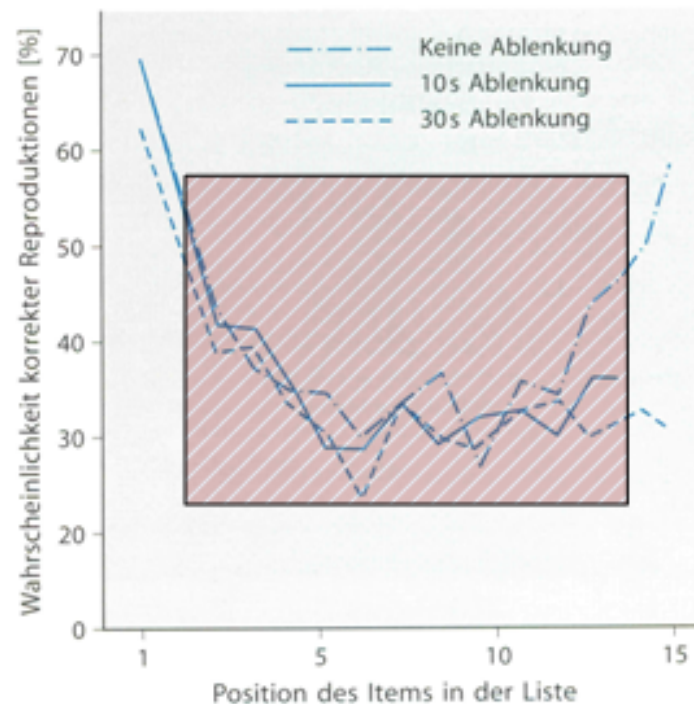


Abbildung 5.1: Durch die Reduzierung auf vier bis sechs Kommandos pro Systemausgabe, kann der Primacy- und Recency-Effekt optimal genutzt werden. Die reproduktionsschwachen Items in der Mitte von Systemausgaben (schraffierte Fläche) werden reduziert bzw. entfernt. Quelle: Zimbardo und Gerrig (1999)

Kommandos sinnvoll, da der Primacy- und Recencyeffekt noch akzeptable Reproduktionswerte zulässt und dennoch die Hilfe nicht in zu viele Ebenen unterteilt wird (Wirth, 2002).

Die Anordnung der Sprachkommandos erfolgt anhand der Verwendungshäufigkeit¹ bzw. dem Komplexitätsgrad bestimmter Funktionen und kommt der Forderung nach, den Einstieg in ein komplexes System anhand einfacher und häufig genutzter Funktionen zu erleichtern (vgl. Kapitel 4.1). Für Dialog 3.8 (12 Kommandos) würde dies bedeuten, dass dieser nach einer Aufteilung aus drei Dialogschritten bestünde² (siehe Dialog 5.2)

Die Ausgabe von Hilfeinhalten geordnet nach Anwendungshäufigkeit und Komplexität trägt anfänglich dazu bei, Benutzern einen leichteren Einstieg in den Um-

¹Die Verwendungshäufigkeit einzelner Funktionen geht aus internen Studien über den Umgang von 5er BMW Fahrern mit dem iDrive-Bediensystem hervor (Gründl, 2006).

²Da diese Arbeit in Zusammenarbeit mit der BMW Group entstand und die Ergebnisse in die Weiterentwicklung des iDrive Nachfolgers einfließen, erfolgen die nachfolgenden Betrachtungen praktischer Natur stets mit Fokus auf das iDrive SDS.

User:	Optionen.
System:	Optionen. Nennen Sie einen Namen gefolgt von anrufen, z.B. Markus Abel anrufen, oder sagen Sie Name wählen, Nummer wählen oder Telefonbuch. Um weitere Kommandos zu hören, sagen Sie weiter.
User:	Weiter.
System:	Weitere Kommandos sind Wahlwiederholung, entgangene Anrufe, Top acht oder Nummer speichern. Um weitere Kommandos zu hören, sagen Sie weiter.
User:	Weiter.
System:	Weitere Kommandos sind Kurzwahlliste, angenommene Anrufe, aktive Anrufe und Telefon aus.

Dialog 5.2: Strukturierter Hilfedialog im Kontext Telefon

gang mit einem Bedienkonzept zu ermöglichen. Mit zunehmender Erfahrung in verschiedenen Anwendungsbereichen werden jedoch entsprechende Informationen in der Hilfe irrelevant. Da der Kompetenzerwerb bei jedem Benutzer verschieden stark ausgeprägt ist und die thematischen Schwerpunkte variieren, scheint die individuelle Anpassung sowohl der Hilfeinhalte als auch der Dialogstrategie an gesammelte Erfahrungen und Präferenzen sinnvoll.

5.2.1 Antizipative Reorganisation der Hilfeinhalte durch Analyse des Benutzerverhaltens

Im Hinblick auf eine optimale Hilfefunktionalität ist es von entscheidender Bedeutung, ein Modell über das Wissen des Benutzers zu etablieren. Beim Erstkontakt mit dem SDS hat der Benutzer noch kein Modell über die Systemfunktionalität ausgebildet (siehe Abbildung 5.2). Ziel beim Erstkontakt muss es also sein,

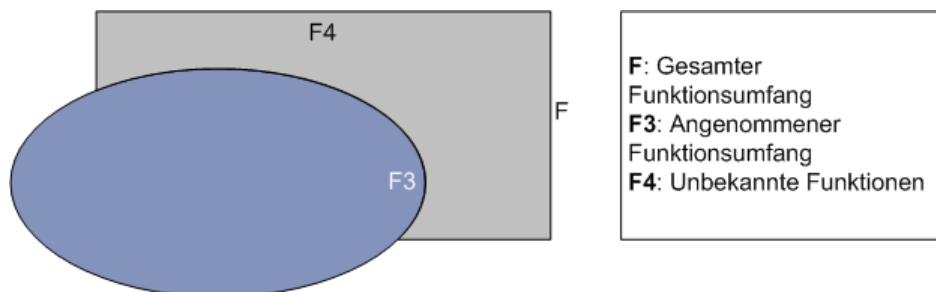


Abbildung 5.2: Beim Erstkontakt eines Benutzers mit dem SDS existiert kein Modell über die Funktionalität des Systems. Lediglich die Menge F3 beschreibt Annahmen des Benutzers, welche Funktionalität das System beinhalten könnte.

die Funktionsumfänge F1 und F2 (siehe Kapitel 4.1) aus F zu etablieren. Dazu werden in der Hilfe zunächst grundlegende Systemfunktionen erläutert. Sobald

Ebene 1	p	Ebene 2	p	Ebene 3	p
Kommando A	1	Kommando E	5	Kommando I	9
Kommando B	2	Kommando F	6	Kommando J	10
Kommando C	3	Kommando G	7	Kommando K	11
Kommando D	4	Kommando H	8	Kommando L	12

Abbildung 5.3: Abstrakte Struktur des Hilfesystems beim Erstkontakt mit dem SDS: grundlegende und allgemein häufig verwendete Funktionen (Kommandos A bis F) stehen am Anfang der Hilfeausgaben, komplexe und allgemein selten verwendete Funktionen (Kommandos G bis L) am Ende einer Ausgabe.

sich ein Funktionsumfang F1 herausbildet, d.h. Funktionen häufig verwendet werden und als gelernt betrachtet werden können, können diese Funktionen aus der ersten Hilfeebene entfernt werden, da diese keinen Mehrwert bedeuten. Vielmehr muss die Hilfe Funktionen abdecken, die ein Benutzer seltener verwendet und die ihm weniger geläufig sind (F2). Zudem soll er einen Überblick über die Möglichkeiten des Systems erhalten (F3 innerhalb F) und Hilfe zu Funktionen erhalten, die er noch nicht kennt (F4). Angenommen ein Benutzer erlernt durch häufige Anwendung die Sprachkommandos A und C aus Abbildung 5.3, können diese aus der ersten Hilfeebene entfernt und in die letzte Ebene verschoben werden (siehe Abbildung 5.4).

Um das Wissen eines Benutzers über einzelne Funktionen bestimmen zu können, ist eine Analyse des individuellen Benutzungs- und Lernverhaltens über einen längeren Zeitraum notwendig. Wird eine Funktion häufig verwendet, kann angenommen werden, dass diese durch Übung nach einer bestimmten Zeitspanne erlernt wird (F1). Das Wissen über diese Funktion bleibt jedoch nur über einen begrenzten Zeitraum erhalten und gerät nach einiger Zeit der Nichtnutzung wieder in Vergessenheit. Somit muss die entsprechende Funktion wieder dem Funktionsraum F2, F3 oder F4 zugeordnet und in der Hilfe angeboten werden. Beide Vorgänge, sowohl das Lernen als auch das Vergessen, sind hochgradig dynamisch und individuell sehr verschieden. Im Bereich der Human Factors-Forschung und der kognitiven Neuropsychologie existieren jedoch Modelle, welche die komplexen Prozesse des Lernens und Vergessens mathematisch beschreiben. Anhand dieser Modelle ist es möglich, das Wissen eines Benutzers über ein System zu ermitteln und für die Anpassung des Systems zu verwenden.

Ebene 1	p	Ebene 2	p	Ebene 3	p
Kommando B	1	Kommando G	5	Kommando K	9
Kommando D	2	Kommando H	6	Kommando L	10
Kommando E	3	Kommando I	7	Kommando A	11
Kommando F	4	Kommando J	8	Kommando C	12

Abbildung 5.4: Veränderte Anordnung der Sprachkommandos: Unter der Annahme, dass die Kommandos A und C gelernt wurden, werden diese an Position $p = 11$ und $p = 12$ in der letzten Hilfeebene verschoben. Die restlichen rücken in der Liste um zwei Positionen nach oben.

Dabei muss jedoch darauf geachtet werden, dass die Anpassung nicht zu Irritation seitens des Benutzers führt. Dies ist ein grundlegendes Problem adaptiver Systeme, da etablierte Modelle über das Systemverhalten unter Umständen nicht mehr gültig sind und somit die erwartungskonforme Verhalten des Systems nicht mehr gegeben ist (Krause, 1988; Benyon und Murray, 1993). Um etwaigen negativen Einflüssen vorzubeugen, muss das Konzept der Adaption zum einen explizit in der allgemeinen Hilfe erläutert werden und zum anderen technisch eher konservativ ausgelegt sein, d.h. die Adaption darf nicht zu stark ausgeprägt sein und soll nach Möglichkeit vom Benutzer nicht bemerkt werden. Ferner muss der Benutzer die Möglichkeit haben, ein verändertes System wieder in den Ursprungszustand zurückzusetzen, falls die Adaption des Systems unzufriedenstellend ausfällt (Johansson, 2004).

5.2.1.1 Beschreibung des Lernens anhand individueller Lernkurven

Der Vorgang des Lernens bestimmter Fähigkeiten ist ein komplexer kognitiver Prozess, für den verschiedene Erklärungsansätze existieren. Ungeachtet der jeweiligen Theorie existieren verschiedene Modelle, die den Prozeß des Erlernens komplexer Fähigkeiten mathematisch beschreiben.

In Totzke et. al. (2003b) wird speziell für die Domäne der visuell haptischen Anzeige-Bedienkonzepte im Fahrzeug die Verwendung des *Potenzgesetzes der Übung* nach Newell und Rosenbloom (1981) vorgeschlagen. Das Potenzgesetz der Übung erlaubt grundsätzlich Aussagen darüber, in welcher Zeit eine bestimmte Aufgabe gelöst wird. Mit zunehmender Übung nimmt die zur Lösung benötigte Zeit kontinuierlich ab (siehe Abbildung 5.5). Nach Anderson (2001) geht die zur Lösung einer Aufgabe benötigte Zeit der kognitiven Komponente mit zunehmen-

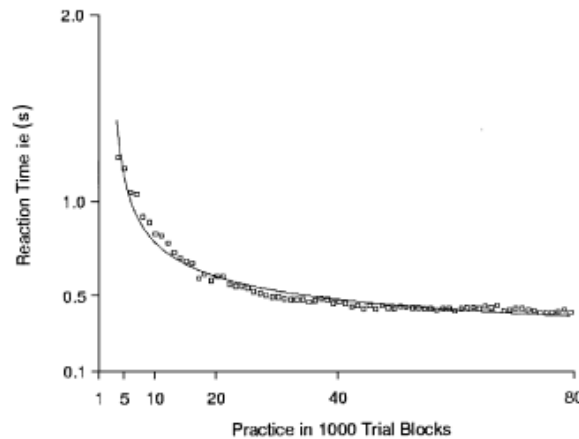


Abbildung 5.5: Exemplarischer Verlauf einer Lernkurve: Zeit die benötigt wird, um eine einfache Aufgabe auszuführen. Quelle: Ritter und Schooler (2002)

der Übung gegen Null. Zugleich impliziert das Potenzgesetz, dass der Lernprozess anfänglich mit großen Lernerfolgen verbunden ist, mit zunehmender Übung die Verbesserung der Lernleistung jedoch relativ abnimmt. Das Potenzgesetz der Übung kann nach Ritter und Schooler (2002) durch folgende Gleichung beschrieben werden:

$$T = B \cdot N^{-\alpha} \quad (5.1)$$

Die Bearbeitungszeit T errechnet sich aus den Parametern B für die Zeit beim ersten Bearbeitungsversuch, N für die Anzahl der Wiederholungen und α für den Lernfaktor (Steigung der Funktion).

Überträgt man diese allgemeinen Überlegungen auf die Erlernbarkeit von Sprachkommandos für ein SDS, kann auf diese Weise der Zeitraum bis zum Erreichen eines bestimmten Dialogziels durch die Äußerung eines Sprachkommandos berechnet werden. Von größerem Interesse ist jedoch die Frage, nach wieviel Übung ein bestimmtes Sprachkommando als gelernt beurteilt werden kann bzw. nach wie vielen Wiederholungen N ein Sprachkommando in einer minimalen Zeit T geäußert werden kann. Dazu kann Gleichung 5.1 nach N aufgelöst werden:

$$N = \sqrt[-\alpha]{\frac{T}{B}} \quad (5.2)$$

Die Zeit T beschreibt nun den Zeitraum, der im Optimalfall zum Erreichen eines Dialogziels benötigt wird. Da die kognitive Komponente für die Ausführung eines Sprachkommandos bei ausreichender Übung gegen Null tendiert, wird T einzig durch die motorische Komponente der Aktivierung des SDS und durch die Zeit für die Äußerung eines Sprachkommandos bestimmt. Die minimale Zeit, die für

diese Aktionen notwendig ist, beträgt für das iDrive System zwei Sekunden. Der Lernkoeffizient α ist die einzige nicht messbare Größe in dieser Gleichung und muss durch experimentelle Untersuchungen bestimmt werden (siehe Kapitel 5.4.1).

Das Potenzgesetz der Übung dominiert seit dessen Veröffentlichung die Beschreibungsmodelle in der Lernforschung, Heathcote et. al. (2002) kritisieren jedoch die dem Potenzgesetz zugrunde liegende Annahme, der Lernprozess verlöre mit zunehmender Übung an Effizienz. Vielmehr gehen sie davon aus, Lernen sei ein konstanter Prozess und könne folglich mathematisch durch eine Exponentialfunktion beschrieben werden:

$$T = B \cdot e^{-\alpha \cdot N} \quad (5.3)$$

Die Parameter sind identisch zu denen des Potenzgesetzes aus Gleichung 5.1. Löst man Gleichung 5.3 ebenfalls nach N auf, so erhält man:

$$N = \frac{\ln T - \ln B}{-\alpha} \quad (5.4)$$

Welches der beiden Modelle für die Domäne dieser Arbeit zutreffend ist, wird in Kapitel 5.4.1 experimentell bestimmt.

Entscheidend ist jedoch, dass N einen Schwellenwert markiert, anhand dessen eine Aussage darüber getroffen werden kann, nach wie vielen Wiederholungen eine Funktion als gelernt beurteilt werden kann.

Um die Hilfeinhalte dynamisch anzuordnen, erhält jedes Sprachkommando eine Positionsangabe p sowie einen Laufindex i (siehe Abbildung 5.6); p beschreibt die Position des Sprachkommandos im Hilfedialog, i beschreibt die Verwendungshäufigkeit bzw. die Wiederholungen eines Sprachkommandos. Mit jeder Verwendung eines Sprachkommandos erhöht sich dessen Index i um den Wert 1 ($i_{neu} = i_{alt} + 1$). Ist der Index i_F einer Funktion F kleiner N_F , bedeutet dies, dass die Funktion

Ebene 1	p	i	N	Ebene 2	p	i	N	Ebene 3	p	i	N
Kommando A	1	0	0	Kommando E	5	0	0	Kommando I	9	0	0
Kommando B	2	0	0	Kommando F	6	0	0	Kommando J	10	0	0
Kommando C	3	0	0	Kommando G	7	0	0	Kommando K	11	0	0
Kommando D	4	0	0	Kommando H	8	0	0	Kommando L	12	0	0

Abbildung 5.6: Parameter der Funktionen für die Hilfeausgabe

noch nicht oft genug angewendet wurde, um als gelernt angesehen zu werden. Im Gegensatz dazu gilt eine Funktion als gelernt, wenn i_F größer als N_F ist.

Neben direkten Lernstrategien (Wiederholen von Sprachkommandos) können Benutzer innerhalb eines Bediensystems auch Wissen durch Analogiebildung über ähnliche Funktionen erzeugen. Beherrscht ein Benutzer beispielsweise das CD Kommando „Nächster Titel“, kann angenommen werden, dass er Kommandos einer Funktionsfamilie ϕ (z.B. „Vorheriger Titel“) ableiten kann und dieses Wissen zudem auf den Bereich der DVD Bedienung übertragen kann. Falls also ein Kommando verwendet wird, das Bestandteil einer Funktionsfamilie ist, muss sich diese Form des Lernens durch eine Erhöhung um den Wert γ auf die Indices der einzelnen Funktionen auswirken:

$$i_{F_{alt}} = i_{F_{neu}} + \gamma \quad (5.5)$$

Die experimentelle Bestimmung des Werts γ erfolgt in Kapitel 5.4.2.

5.2.1.2 Beschreibung des Vergessens anhand individueller Vergessenskurven

Die Modellbildung im Bereich des Vergessens geht zurück auf die Vergessenskurve nach Ebbinghaus. In Ebbinghaus (1885) werden Ergebnisse von Gedächtnisexperimenten vorgestellt, bei denen Probanden eine Abfolge von Nonsenssilben lernen und nach einem bestimmten Zeitraum reproduzieren mussten. Als Erfolgsmaß zog Ebbinghaus den Anteil an korrekt reproduzierten Silben heran. Überträgt man die Ergebnisse der Experimente in ein Diagramm, so ergibt sich das charakteristische Bild der Vergessenskurve (siehe Abbildung 5.7).

Anfänglich fällt die Vergessenskurve steil ab, d.h. innerhalb kurzer Zeit werden relativ viel Informationen vergessen. Mit zunehmender Zeit wird die Kurve jedoch flacher und tendiert letztendlich gegen Null. Dieser Verlauf der Vergessenskurve konnte in verschiedenen Experimenten beobachtet werden (Davis et. al., 2003; White, 2003).

Obgleich der Verlauf der Vergessenskurve stets ähnlich ist, existieren zahlreiche mathematische Beschreibungsmodelle. In Rubin und Wenzel (1996) werden 105 Modelle analysiert und mit den Ergebnissen verschiedener Gedächtnisexperimente abgeglichen. Die Untersuchung lieferte zwar keine allgemein anwendbare mathematische Gleichung für das Vergessen, einige Modelle konnten jedoch erfolgreich auf eine größere Anzahl von Studien angewendet werden. Dazu zählen eine logarithmische Funktion (5.6), eine Exponentialfunktion (5.7), eine Potenzfunktion

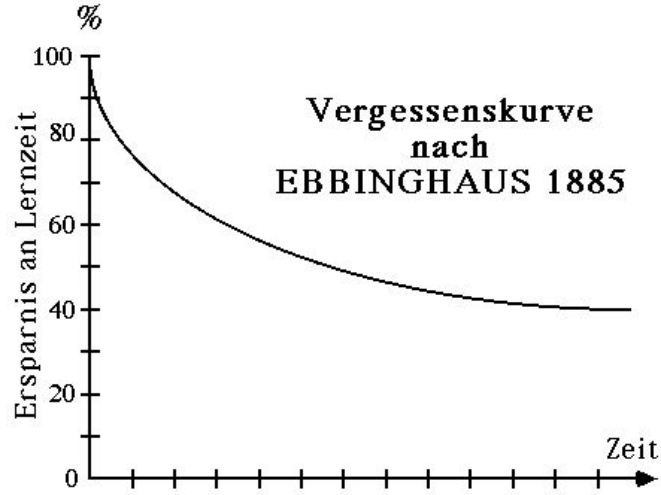


Abbildung 5.7: Exemplarischer Verlauf einer Vergessenskurve. Quelle: Stangl (1997)

(5.8) und eine weitere Exponentialfunktion mit Wurzel (5.9).

$$\mu_{neu} = \mu_{alt} \cdot \ln T + e^{-\delta} \quad (5.6)$$

$$\mu_{neu} = \mu_{alt} \cdot e^{-\delta \cdot T} \quad (5.7)$$

$$\mu_{neu} = \mu_{alt} \cdot (T + \delta)^{-\delta} \quad (5.8)$$

$$\mu_{neu} = \mu_{alt} \cdot e^{-\delta \cdot \sqrt{T}} \quad (5.9)$$

Die Variable μ_{alt} repräsentiert die Menge ursprünglich gelernter Informationen und μ_{alt} beschreibt die Menge an Informationen, die nach einem Zeitraum T noch erinnert werden kann. Der Vergessenskoeffizient δ beschreibt die Steigung der Vergessenskurve und ist somit ein Maß für die Geschwindigkeit des Vergessens. Welche dieser Gleichungen den Vergessenprozess im Umgang mit dem iDrive SDS am besten beschreibt, muss analog zu den Lernkurven experimentell ermittelt werden (vgl. Kapitel 5.4.3).

Die Integration des Vergessenprozesses in das Hilfesystem erfolgt über den Index i_F einer Funktion. Nach n Anwendungen einer Funktion erhöht sich i_F um diesen Wert n . Wird diese Funktion einige Zeit nicht verwendet, nimmt i_F gemäß einer Vergessensfunktion $f(T)$ (logarithmische Funktion, Exponentialfunktion, Potenzfunktion oder Exponentialfunktion mit Wurzel) wieder ab (siehe Abbildung 5.8):

$$i_{F_{neu}} = i_{F_{alt}} \cdot f(T) \quad (5.10)$$

Je nachdem, ob $i_F > N$ oder $i_F < N$ ist, gilt eine Funktion F als gelernt oder nicht gelernt. Dialog 5.3 soll verdeutlichen, wie sich i_F auf die Hilfe auswirkt.

User:	Optionen.	<i>Hilfeausgaben in der ursprünglich festgelegten Reihenfolge</i>
System:	Optionen. Sagen Sie Maßstab in Meter oder Kilometer, Karte nordweisend, Karte fahrtweisend, Perspektive oder Routenkriterien. Um weitere Sprachkommandos zu hören, sagen Sie weiter.	
User:	Weiter.	
System:	Sagen Sie Sprachhinweise an, Sprachhinweise aus, Zielführung an oder Zielführung aus.	
User:	\langle mehrmalige Verwendung des Kommandos „Perspektive“ \rangle	$i_{Perspektive} > N,$ $\Rightarrow pos_{Perspektive} = 10$
User:	Optionen.	
System:	Optionen. Sagen Sie Maßstab in Meter oder Kilometer, Karte nordweisend, Karte fahrtweisend, Routenkriterien oder Sprachhinweise an. Um weitere Sprachkommandos zu hören, sagen Sie weiter.	
User:	Weiter.	
System:	Sagen Sie Sprachhinweise aus, Zielführung an, Zielführung aus oder Perspektive.	
		<i>Da „Perspektive“ gelernt wurde, befindet es sich nun am Ende der Hilfeausgabe.</i>

Dialog 5.3: Beispieldialog.

5.2.2 Generische Anpassung der Dialogstrategie an Nutzererfahrung

Neben der Anpassung der Hilfeinhalte scheint es sinnvoll, die Dialogstrategie des SDS an die Bedürfnisse eines Benutzer anzupassen. Dazu können verschiedene Herangehensweisen in Betracht gezogen werden.

Komatani et. al. (2003) entwickelten ein sprachbasiertes Auskunftssystem für den Personennahverkehr, das sich sowohl an Fähigkeiten und Wissen des Benutzers als auch und an dessen Zeitdruck anpasst. Die Etablierung eines Benutzermodells basiert auf der Auswertung der Dialoginteraktionen. Die Fähigkeiten eines Benutzers werden anhand von Faktoren wie der Anzahl an benutzerinitiierten Interaktionen, der Ausführlichkeit der Systemanfragen oder der Anzahl der Wiederholung einzelner Dialogschritte gemessen. Das Wissen des Benutzers leitet sich beispielsweise aus genauen Ortsangaben des Benutzers ab, was darauf schließen lässt, dass dieser in gewissem Umfang ortskundig ist. Der Zeitdruck wird aus der Güte der Spracheingaben (Spracherekennerergebnis) und der Verwendung von

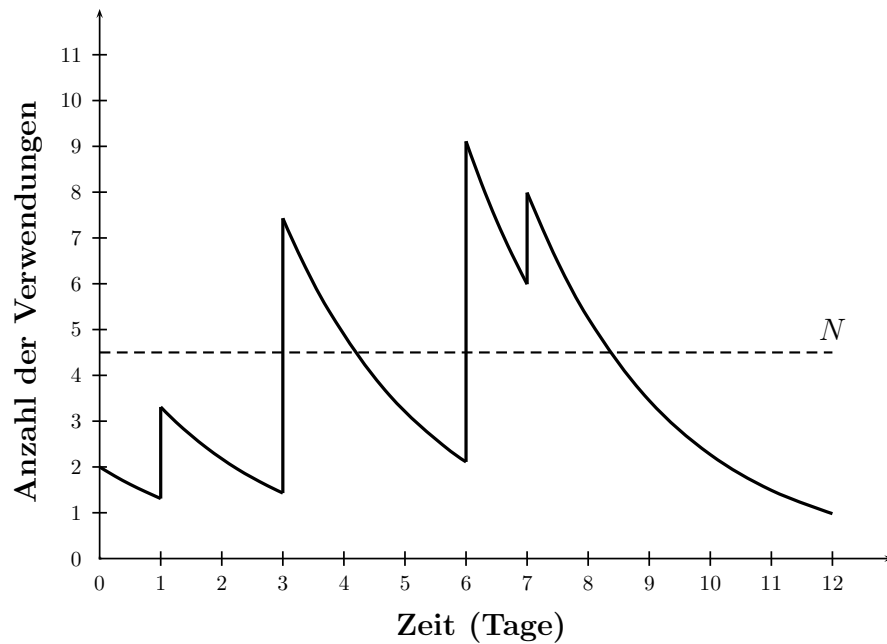


Abbildung 5.8: Exemplarischer Verlauf von i_A in Anlehnung an Abbildung 5.6: Mit n Verwendungen pro Tag steigt der Index der Funktion A um den Wert n an. Solange die Funktion nicht verwendet wird, sinkt i_A gemäß der Vergessenskurve ab. Solange $i_A < N$, bleibt Kommando A in der Liste der Hilfe auf Position 1. Wird $i_A > N$, erhält Kommando A Position 12 in der Hilfeausgabe, die restlichen Kommandos rücken um eine Position nach oben.

Barge-In ermittelt. Anhand dieser Aspekte erfolgt eine Anpassung der Dialogstrategie des SDS, d.h. die Systemausgaben können zwischen detaillierten und ausführlichen und kurzen Informationsdialogen variieren.

Auch Bohnenberger (2001) stellte Untersuchungen mit dem Ziel an, durch die Berücksichtigung verschiedener Parameter in den sprachlichen Äußerungen Dialogstrategien für die Ausgabe von Handlungsanleitungen zu entwickeln. Wird beispielsweise eine erhöhte kognitive Belastung anhand eindeutiger sprachlicher Merkmale registriert, erfolgen Handlungsanweisungen sequentiell in kleinen informationellen Einheiten. In einer gegenteiligen Situation werden Informationen hingegen kumuliert ausgegeben.

Einen ähnlichen Ansatz verfolgten Müller et. al. (2001) bei der Entwicklung eines mobilen Auskunftssystems für einen Flughafen. Die Adaption der Sprachdialoge wirkte sich auf die Ausführlichkeit der Systemausgaben aus. Der Nachteil dieser Systeme ist jedoch, dass die Adaption nur kurzzeitig erfolgt und kein langfristiges Modell über den Erfahrungsgrad eines Benutzers ausbildet. In automotiven Umgebungen ist gerade dies ein wichtiges Kriterium.

Hassel (2006) verwendet daher einen Adaptionsmechanismus, welcher die Dialogstrategie längerfristig an die Erfahrungen eines Benutzers anpasst. Die Gene-

rierung eines individuellen Benutzermodells erfolgt nach Hassel anhand der Parameter Hilfeanfragen h , Optionenanfragen o , Timeouts t , ASR-Failures e und Onset Time T_m . Sinnvoll erscheint eine Erweiterung der Interaktionsparameter um Barge-In b (siehe Tabelle 5.1). Alle Parameter können implizit aus der regulären Interaktion des Benutzers mit dem SDS extrahiert werden. Da diese Form der Adaption bereits erfolgreich an einem prototypischen SDS im Fahrzeug getestet wurde, können die bestehenden Konzepte und Algorithmen übernommen und erweitert werden.

Parameter π	Bedeutung
Hilfeanfragen, h	Benutzer fragt nach genereller Information über die Systembedienung.
Optionsanfragen, o	Benutzer fragt nach Sprachkommandos im aktuellen Kontext.
Timeouts, t	Der Spracherkenner erhält kein akustisches Signal.
ASR-Failure, e	Der Spracherkenner kann das akustische Signal nicht verarbeiten. Das kann an OOV-Eingabe oder undeutlicher Aussprache liegen.
Onset Time, T_m	Zeit, die der Benutzer braucht, um mit seiner Eingabe zu beginnen.
Barge-In, b	Der Benutzer beginnt mit der Spracheingabe, bevor die Systemausgabe beendet ist.

Tabelle 5.1: Interaktionsparameter zur Bestimmung eines Benutzermodells aus Hassel (2006), erweitert um den Parameter Barge-In.

Jeder Dialog wird anhand der Interaktionsparameter bewertet und das User Model \vec{UM} wird für jede Funktion φ durch einen Vektor \vec{UM}_φ abgebildet (siehe Dialog 5.4). Die Gewichtung der einzelnen Interaktionsparameter erfolgt anhand eines Gewichtsvektors \vec{UM}_G .

Der Gewichtsvektor \vec{UM}_G drückt die Aussagekraft jeder Komponente des Benutzmodells \vec{UM} hinsichtlich der Nutzererfahrung mit dem System aus (Hassel, 2006).

Die Gewichtungen der einzelnen Komponenten v_{G_π} von \vec{UM}_G müssen in Anlehnung an die Multi-Attribute-Utility-Theorie (MAUT) nach Schütz und Schäfer (2002) der Bedingung

$$\sum_{i=1}^n v_{G_\pi} = 1$$

entsprechen, d.h. es muss gelten:

$$v_{G_h} + v_{G_o} + v_{G_t} + v_{G_e} + v_{G_{T_m}} + v_{G_b} = 1$$

$h = 0, o = 0, t = 0, e = 0, T_m = 0, b = 0$		
User:	$\langle PTT \rangle$	
System:	Spracheingabe.	
User:	$\langle Unverständliches \rangle$	(1) $e = e + 1$
System:	Wie bitte?	
User:	$\langle Unverständliches \rangle$	(2) $e = e + 1,$ $T_m = T_m + 1$
System:	Ich konnte Sie nicht verstehen. Wählen Sie eine Frequenz, Station oder einen Sender.	
User:	$\langle Unverständliches \rangle$	(3) $e = e + 1$
System:	Ich konnte Sie nicht verstehen. Spracheingabe beendet.	
User:	$\langle PTT \rangle$	
System:	Spracheingabe.	
User:	$\langle Pause \rangle$	(4) $t = t + 1$
System:	Ich konnte Sie nicht verstehen. Bitte wiederholen Sie.	
User:	Optionen.	(5) $o = o + 1$
System:	Optionen. Wählen Sie eine Frequenz, Station oder einen Sender...	
User:	Sender wählen.	(6) $T_m = T_m + 1,$ $b = b + 1$
$\vec{UM}_{\text{Frequenz wählen}} = [h, o, t, e, T_m, b] = [0, 1, 1, 3, 2, 1]$		

Dialog 5.4: Beispieldialog für die Berechnung von \vec{UM} für die Funktion „Frequenz wählen“.

Da es aus technischen Gründen nicht möglich war, die Onset Time T_m zu bestimmen, muss diese aus der Berechnung des User Models ausgenommen werden. Die Werte für v_{G_π} ermittelte Hassel (2006) durch eine Bestimmung der Produkt-Moment-Korrelation nach Pearson³ für die Parameter Anwenderklassifikation, Hilfe- und Optionsanfragen, Timeouts und ASR-Fehler anhand von 229 Dialogen mit dem iDrive SDS. Die Ergebnisse sind in Tabelle 5.2 ersichtlich. Der stärkste Zusammenhang besteht zwischen dem Erfahrungsgrad und den ASR-Fehlern. Dies kann damit begründet werden, dass unerfahrene Benutzer häufig Vokabular verwenden, das nicht in der Grammatik des Spracherkenners enthalten ist.

³Die Produkt-Moment-Korrelation nach Pearson $r_{x,y}$ berechnet sich nach folgender Formel (Rasch et. al., 2004):

$$r_{x,y} = \frac{\widehat{cov}(x,y)}{\widehat{\sigma}_x \cdot \widehat{\sigma}_y} = \frac{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2} \cdot \sqrt{\frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2}}$$

Im Zähler steht die Kovarianz der beiden Variablen x und y , im Nenner stehen die Standardabweichungen beider Variablen.

Die daraus resultierenden Fehlerkennungen werden der Kategorie der ASR-Fehler zugeordnet.

Die Korrelation zwischen dem Erfahrungsgrad und der Verwendung von Barge-In wurde nicht empirisch erhoben, weshalb für den Umfang dieser Arbeit ein Korrelationswert von $r_{b,AE} = 0,30$ angenommen wird⁴. Die Verwendung von Barge-In ist ein Indiz für einen Experten und trägt somit im Gegensatz zu den anderen Komponenten positiv zur Dialogqualität bei.

	Anfänger- bzw. Expertenverhalten
Hilfeanfragen h	$r_{h,AE} = -0,188$
Optionsanfragen o	$r_{o,AE} = -0,241$
Timeouts t	$r_{t,AE} = -0,175$
ASR-Fehler e	$r_{e,AE} = -0,498$

Tabelle 5.2: Produkt-Moment-Korrelation nach Pearson: das negative Vorzeichen drückt die negative Beziehung zwischen den Komponenten und des beobachteten Erfahrungsgrads aus, d.h. diese Komponenten verschlechtern die Dialogqualität. Quelle: Hassel (2006)

Als Anhaltspunkt für die Stärke des Zusammenhangs kann folgende Kategorisierung dienen:

$r_{x,y} = 0,00 $	→ keine Korrelation
$ 0,00 < r_{x,y} \leq 0,20 $	→ sehr schwache Korrelation
$ 0,20 < r_{x,y} \leq 0,40 $	→ schwache Korrelation
$ 0,40 < r_{x,y} \leq 0,60 $	→ mittlere Korrelation
$ 0,60 < r_{x,y} \leq 0,80 $	→ starke Korrelation
$ 0,80 < r_{x,y} < 1,00 $	→ starke Korrelation
$r_{x,y} = 1,00 $	→ perfekte Korrelation

Basierend auf den ermittelten Korrelationswerten erhalten die Parameter v_{G_π} unter der Voraussetzung $\sum_{i=1}^n v_{G_\pi} = 1$ folgende Werte:

$$\begin{aligned}
 -0,188 \cdot x - 0,241 \cdot x - 0,175 \cdot x - 0,498 \cdot x + 0,250 \cdot x &= 1 \\
 -0,802 \cdot x &= 1 \\
 x &= -1,247
 \end{aligned}$$

$$\Rightarrow v_{G_h} = 0,23; v_{G_o} = 0,30; v_{G_t} = 0,22; v_{G_e} = 0,62; v_{G_b} = -0,37;$$

⁴Die Annahme dieses Wertes beruht auf Beobachtungen von Benutzern im Umgang mit dem iDrive SDS. Barge-In wird zwar tendenziell von erfahreneren Benutzern verwendet, jedoch scheint die Korrelation nicht so ausgeprägt zu sein, wie bei ASR-Fehlern und unerfahrenen Benutzern.

Somit ergibt sich folgender Gewichtsvektor für das User Model:

$$\overrightarrow{UM}_G = \begin{pmatrix} v_{G_h} = 0,23 \\ v_{G_o} = 0,30 \\ v_{G_t} = 0,22 \\ v_{G_e} = 0,62 \\ v_{G_b} = -0,37 \end{pmatrix}$$

Auffällig ist, dass die Gewichtung der Erkennenfehler mit $v_{G_e} = 0,62$ sehr stark ausgeprägt ist, d.h. dass eine große Anzahl an Erkennenfehlern ein starkes Indiz für einen unerfahrenen Benutzer darstellt. Dies kann von mangelnder Kenntnis von Sprachbefehlen oder Unsicherheit im Umgang mit dem SDS herrühren. Die Gewichtung von v_{G_b} fällt als einzige negativ aus, da sich der Parameter Barge-In im Gegensatz zu allen anderen Parametern nicht negativ auf die Einstufung des Benutzers als Experte auswirkt. Die Verwendung von Barge-In ist ein Merkmal für einen erfahrenen Benutzer oder Experten.

Unter Berücksichtigung eines Vektors $\overrightarrow{UM}_\varphi$ und dem Gewichtsvektor \overrightarrow{UM}_G kann nun der Vergleichsvektor Δ_{UM} aus dem Skalarprodukt der beiden Vektoren gebildet werden. Basierend auf Dialog 5.4 ergibt sich folgender Vergleichsvektor:

$$\Delta_{UM} = \begin{pmatrix} v_{G_h} = 0,23 \\ v_{G_o} = 0,30 \\ v_{G_t} = 0,22 \\ v_{G_e} = 0,62 \\ v_{G_b} = -0,37 \end{pmatrix} \bullet \underbrace{\begin{pmatrix} h = 0 \\ o = 1 \\ t = 1 \\ e = 3 \\ b = 1 \end{pmatrix}}_{\text{Frequenz wählen}} = 2,01$$

Die Kategorisierung eines Benutzers als Anfänger oder Experte geschieht durch einen Vergleich des Vergleichsvektors mit einem Schwellenwert σ . Ist der Vergleichsvektor größer als σ , wird der Benutzer als Anfänger eingestuft. Ist er kleiner als σ , gilt der Benutzer als Experte. Der Schwellenwert errechnet sich nach Hassel (2006) aus dem Dreifachen des kleinsten Gewichtungsparameters, somit gilt:

$$\sigma = v_{G_t} \cdot 3 = 0,66$$

Um sicherzustellen, dass keine zufällige Zuordnung in eine Kategorie erfolgt, muss ein Benutzer den Schwellenwert drei mal in Folge unter- oder überschreiten. Die Kategorisierung des Benutzers kann dabei in verschiedenen Menüs unterschiedlich ausfallen, z.B. kann jemand im Bereich Navigation als Experte eingestuft werden, während er in den restlichen Bereichen noch den Status eines Anfängers besitzt.

Dieses User Model kann dazu verwendet werden, die Dialogstrategie für das Hilfesystem an den Erfahrungsgrad des Benutzers im Umgang mit dem SDS anzupassen. Unerfahrene Benutzer erhalten eine Hilfe, deren Fokus auf der Wissens-

vermittlung liegt. Die Prompts besitzen daher einen erläuternden Charakter und beinhalten Beispiele zur Verwendung von Funktionen. Um die Systemausgaben dennoch so kurz wie möglich zu halten, enthalten die Hilfedialoge nur vier Items pro Dialogschritt (siehe Dialog 5.5).

User:	Optionen.
System:	Optionen. Sagen Sie Maßstab wählen oder Maßstab in Meter oder Kilometer, z.B. Maßstab 5 Kilometer, Ansichten, Zielführung an oder Sprachhinweise an. Um weitere Sprachkommandos zu hören sagen Sie weiter.
User:	Weiter.
System:	Sagen Sie Routenkriterien, Sprachhinweise aus oder Zielführung aus.

Dialog 5.5: Anpassung der Dialogstrategie im Kontext Navigation/Karte: Unerfahrene Benutzer erhalten eine Hilfe mit vier Items pro Dialogschritt und Beispielen für die Anwendung einzelner Kommandos.

Erfahrene Benutzer erhalten eine Hilfe, die hauptsächlich auf die Präsentation von Informationen ausgerichtet ist. Wenn ein Benutzer vertraut ist im Umgang mit dem SDS, dient die Hilfe entweder als Gedächtnisstütze oder dazu, mehr Informationen über weitergehende Funktionalitäten des SDS zu erhalten. Daher beinhaltet die Hilfe für Experten fünf Items pro Dialogschritt, wobei auf die Ausgabe von Beispielen oder die Konstruktion von Satzstrukturen verzichtet wird (siehe Dialog 5.6). Des weiteren entfallen Metakommandos zur Steuerung der Hilfe („Weiter“), da vorausgesetzt wird, dass diese durch die häufige Anwendung des SDS bereits bekannt sind. Zudem wird das Kommando „Weiter“ auch visuell repräsentiert (siehe Kapitel 5.2.3).

User:	Optionen.
System:	Optionen. Maßstab in Meter oder Kilometer, Karte nordweisend, Karte fahrtweisend, Perspektive, Routenkriterien.
User:	Weiter.
System:	Sprachhinweise an, Sprachhinweise aus, Zielführung an, Zielführung aus.

Dialog 5.6: Anpassung der Dialogstrategie im Kontext Navigation/Karte: Erfahrene Benutzer erhalten eine Hilfe mit fünf Items pro Dialogschritt. Die sprechbaren Kommandos werden in Form einer Liste ausgegeben.

5.2.3 Visuelle Repräsentation der allgemeinen und kontextspezifischen Hilfe

In Anlehnung Paivio (1991) erfolgen Lernprozesse effektiver, wenn sprachlich-akustische Informationen mit visuell-räumlichen Informationen verknüpft werden. Der Theorie der dualen Enkodierung zufolge wird bei der Verarbeitung textueller Information zum einen der Wortgehalt eines Begriffs in einem sprachlich-akustischen Gedächtnisspeicher abgelegt. Gleichzeitig wird eine bildhafte Repräsentation des Begriffs im räumlich-visuellen Gedächtnisspeicher enkodiert. Obgleich die Theorie der dualen Enkodierung nicht unumstritten ist, zeigten verschiedenen Experimente, dass Begriffe mit bildhaften Repräsentationen besser im Gedächtnis bleiben als Begriffe ohne Bildvorstellung (Bildüberlegenheitseffekt). Ebenso führte wiederholtes Aussprechen von Wörtern zu keiner signifikant besseren Behaltensleistung (Hasebrook, 1995).

Die Mehrfachkodierung von Wissen spielt nach Totzke et. al. (2003a) eine entscheidende Rolle beim Kompetenzerwerb für Informationssysteme. So werden bestimmte Begriffe eines Systems Oberbegriffen zugeordnet („Die Routenkriterien befinden sich im Navigationsmenü“) oder räumliche Vorstellungsbilder des Systems aufgebaut („Die Navigation befindet sich im rechten Bereich des Systems“). Ferner trägt auch die motorische Komponente zum Lernprozeß bei („Um zum Telefon zu gelangen, muss ich die Menütaste drücken und einmal nach oben schieben“).

Sowohl die semantische Struktur des Systems als auch dessen räumliche und motorischen Aspekte beeinflussen den Lernaufwand und sind für mögliche Lernprobleme mitverantwortlich. Zusätzlich ergibt sich, dass die Förderung einer Mehrfachkodierung mit einer höheren Lernleistung einhergeht. (Totzke et. al., 2003a)

Aus diesem Grund beinhaltet das Hilfesystem neben der sprachlichen auch eine visuelle Komponente. Die visuelle Repräsentation stellt eine ergänzende Form der Informationsausgabe dar, die den Benutzer zum einen beim Erlernen von Sprachkommandos nach dem Konzept der dualen Enkodierung unterstützt (Salmen, 2003). Zum anderen dient sie als Unterstützung, falls der Benutzer die sprachliche Hilfe nicht richtig verstanden hat oder durch die Fahraufgabe kurzzeitig vom Sprachdialog abgelenkt war. Somit muss der Sprachdialog nicht wiederholt werden, sondern ein kurzer Blick genügt, um einen Überblick über die Liste der Sprachkommandos in der Hilfe zu erhalten.

Die visuelle Repräsentation der Hilfe erfolgt anhand einer separaten Hilfetafel, welche die reguläre Bedienoberfläche überdeckt (siehe Abbildung 5.9). Jede Hilfetafel verfügt über ein Textfeld im oberen Bereich, das Kontextinformationen enthält (allgemeine Hilfe oder Optionen). Farblich abgetrennt davon nimmt der Bereich für die Hilfeinhalte den größten Teil der Hilfetafel in Anspruch. Ober-

bzw. unterhalb der Hilfeinhalte befindet sich ein Pfeilsymbol, welches vorhergehende oder nachfolgende Hilfeebenen kennzeichnet. Die visuelle Hilfe ist mit der sprachlichen Hilfe synchronisiert. Während einer Hilfeausgabe werden die aktuell vorgelesenen Sprachkommandos in der Liste der Sprachbefehle durch eine farbige Umrandung hervorgehoben.

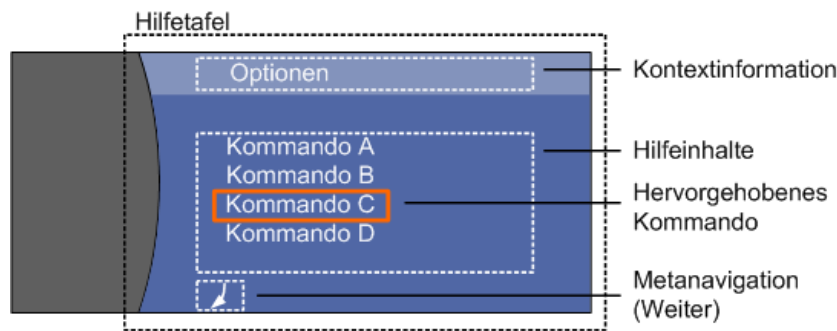


Abbildung 5.9: Schematische Darstellung der Hilfetafel

5.2.3.1 Haptische Bedienung der Hilfe

Die Verwendung eines SDS zur Interaktion mit einem Informationssystem im Fahrzeug birgt viele Vorteile, jedoch besitzt die Spracherkennung als relativ neue Technologie im Fahrzeug noch einige Schwachpunkte. Obwohl die theoretische Erkennrate bei ca. 98% liegt, können auch schlechtere Erkennungsergebnisse zustande kommen. Neben externen Faktoren (z.B. Umgebungsgeräusche bei geöffnetem Fenster) wirken sich v.a. sprecherspezifische Faktoren negativ auf die Erkennrate aus. Spricht ein Benutzer zu laut oder zu leise, mindert dies die Erkennungserleistung ebenso wie verschiedene Akzente oder der sog. Lombard-Effekt (Junqua, 1993). Dieser Effekt tritt häufig nach Fehlerkennungen auf und bewirkt hyperartikulierte Äußerungen. Zu den Symptomen zählen mehr und längere Pausen, eine langsamere und übertrieben deutliche Aussprache sowie eine Veränderung der Intonation. Der Lombard-Effekt ist in der zwischenmenschlichen Kommunikation ein probates Mittel, um Kommunikationsprobleme zu überbrücken. Im Umgang mit Spracherkennungstechnologien führt dieses Verhalten jedoch nur zu einer Verschlechterung des Sprachsignals und somit auch der Interaktionsqualität. Für den Fall einer wiederholt erfolglosen Spracheingabe wechseln Benutzer häufig die Eingabemodalität, um ein Dialogziel zu erreichen (Oviatt, 1999). Studien zeigten, dass multimodale Dialogsysteme die Fehlerrate bei der Sprachbedienung um 19%-41% reduzierten (Oviatt, 2000b).

Aus diesem Grund ist es sinnvoll, neben der Spracheingabe alternative fehlerrobuste Eingabemodalitäten für das Hilfesystem zur Verfügung zu stellen. Geeig-

net ist die Anbindung des Hilfesystems an den Controller, der eine haptische Bedienung zulässt. Durch Drehbewegungen ist es möglich, in der Liste der Hilfefunktionen nach oben oder unten zu navigieren (siehe Abbildung 5.10). Ist das Ende

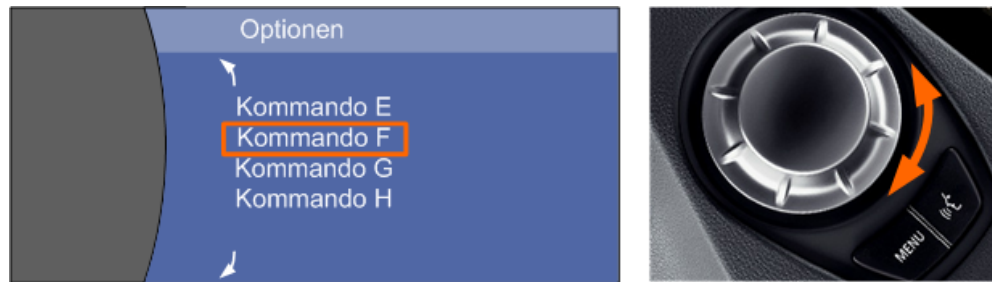


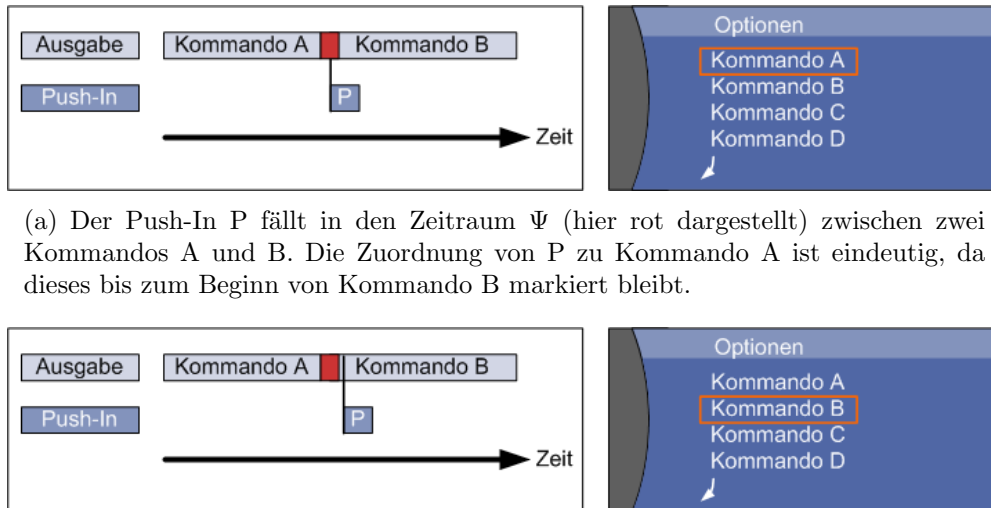
Abbildung 5.10: Visuelle Unterstützung der Hilfe: durch Drehen des Controllers ist es möglich, verschiedene Kommandos anzuwählen und in vorhergehende bzw. nachfolgende Hilfeebenen zu wechseln.

einer Liste erreicht, kann in Analogie zur systemweiten Listennavigation durch eine weitere Drehbewegung in die nächste Hilfeebene gewechselt werden. Diese Aktion stellt das Pendant zum Sprachkommando „Weiter“ dar. Die Metakommandos zur Navigation werden nicht visuell dargestellt, da Versuchsergebnisse in Salmen (2003) aufzeigen, dass sich redundante Metainformationen als verzichtbar erwiesen. Zudem unterstreicht rein sprachliche Ausgabe der Metakommandos den primären Charakter der Sprachbedienung des Hilfesystems. Ferner trägt die reduzierte Informationsdarstellung dazu bei, die zu verarbeitenden Informationen möglichst gering zu halten und die verfügbare Darstellungsfläche optimal für die primären Hilfeinhalte zu verwenden.

Die Auswahl eines Kommandos kann entweder durch Sprache erfolgen oder durch Drücken des Controllers. Spracheingaben sind sowohl am Ende der Systemausgabe als auch während der Sprachausgabe (Barge-In) möglich. In Analogie dazu können mit dem Controller Kommandos nach der Systemausgabe oder währenddessen (Push-In) ausgewählt werden. Gerade die Push-In Funktionalität stellt jedoch eine Herausforderung für das Dialogmanagement dar. Im Gegensatz zum Barge-In ist ein Push-In an einen zeitlich-inhaltlichen Kontext gebunden. Konkret bedeutet dies, dass im Falle eines Push-In analysiert werden muss, welchem Kommando die Auswahl gegolten hat. Besonders im Übergangsbereich zwischen zwei Sprachkommandos muss geprüft werden, ob die Auswahl beispielsweise noch dem letzten vorgelesenen/markierten Kommando zugerechnet werden soll oder schon dem aktuell vorgelesenen/markierten Kommando.

Der wesentliche Faktor bei der Interpretation der Eingabe ist der Zeitraum, in dem diese erfolgt. Kritisch ist eine Zeitspanne Ψ , die zwischen dem Ende eines

vorgesehenen Sprachkommandos und dem Anfang des nächsten Kommandos liegt⁵. Ein Push-In kann genau in die Pause zwischen zwei Kommandos fallen (siehe Abbildung 5.11(a)) oder auch gleich zu Beginn des nächsten Kommandos erfolgen (siehe Abbildung 5.11(b)). Da Kommando A während der entsprechenden Systeme-



(a) Der Push-In P fällt in den Zeitraum Ψ (hier rot dargestellt) zwischen zwei Kommandos A und B. Die Zuordnung von P zu Kommando A ist eindeutig, da dieses bis zum Beginn von Kommando B markiert bleibt.

(b) Der Push-In P erfolgt nach Ψ zu Beginn der Systemausgabe zu Kommando B. In diesem Fall ist die Zuordnung von P zu Kommando A oder B nicht eindeutig und muss durch Methoden der Disambiguierung erfolgen.

Abbildung 5.11: Überlappung multimodaler Ein- und Ausgaben

mausgabe und 70ms danach farblich hervorgehoben bleibt, kann ein Push-In in diesem Fall Kommando A zugerechnet werden. Sobald jedoch die Markierung von Kommando A auf Kommando B wechselt und der Benutzer innerhalb einer Zeitspanne Z einen Push-In durchführt, kann dieser nicht mehr eindeutig Kommando A oder B zugeordnet werden. Es besteht die Möglichkeit, dass der Benutzer nur etwas zu spät auf den Controller gedrückt hat. Um zu klären, ob tatsächlich Kommando B ausgewählt wurde oder der Push-In noch dem Kommando A galt, muss die Zeitspanne Z festgelegt werden und innerhalb Z eine Disambiguierung durchgeführt werden.

Messungen der Reaktionszeit mit Hilfe eines an einen Laptop angeschlossenen Controllers und einem einfachen Programm zur Zeitabnahme zeigten in vergleichbaren Szenarien eine durchschnittliche Reaktionszeit für die Anwendung des Controllers innerhalb von ca. 100ms. Abzüglich der Latenzzeit $\Psi = 70ms$ bleiben für Z 30ms übrig. Push-Ins innerhalb dieser Zeitspanne führen zu einem klärenden Dialogschritt, in dem der Benutzer nochmals zwischen den beiden Kommandos A und B wählen kann (siehe Dialog 5.7). Die Disambiguierungstafel enthält eben-

⁵Für die Implementierung des prototypischen Hilfesystems wurde $\Psi = 70$ Millisekunden gewählt.

System:	Optionen. Sagen Sie Maßstab wählen oder Maßstab in Meter oder Kilometer, z.B. Maßstab 5 Kilometer, Ansichten, Zi...
User:	<i>< Push-In ></i>
System:	Ihre Auswahl war nicht eindeutig. Meinten Sie Ansichten oder Zielführung an?
User:	Ansichten.

Dialog 5.7: Disambiguierungsdialog

falls nur die beiden relevanten Kommandos zur Auswahl. In Analogie zum systemweiten Disambiguierungskonzept sind die beiden Einträge in der Liste nummeriert und können sowohl direkt als auch anhand Ihrer Nummer (z.B. „Eintrag zwei“) aufgerufen werden.

**Abbildung 5.12:** Disambiguierungstafel

5.3 Kontextunabhängige, natürlichsprachliche Hilfe

Die kontextabhängige Hilfe bietet einen guten Überblick über die Kommandos, die in einem bestimmten Kontext sprechbar sind. Dazu muss ein Benutzer diesen Kontext jedoch erst einmal erreichen. Gerade beim Erstkontakt mit einem SDS stellt sich für den Benutzer die Frage, welche Kommandos man sprechen kann und wie gewünschte Dialogziele erreicht werden können. Speziell für diese Situation scheint es sinnvoll, eine Form der Hilfe anzubieten, die unabhängig von einzelnen Kontexten ist und natürlichsprachliche Anfragen des Benutzers verarbeiten kann. Beobachtungen im Rahmen der Arbeit von Hof (2004) ergaben, dass Benutzer, die ein vorgegebenes Dialogziel erreichen sollen und kein mentales Modell über ein SDS oder dessen Hilfesystem besitzen, intuitiv eine Frage formulieren. Diese war entweder an reale Personen (Versuchsleiter) oder als Hilfeanfrage an das SDS gerichtet.

Die Art der Hilfeanfragen lässt sich primär zwei Kategorien einteilen; je nach Dialogziel formulierten die Benutzer Fragen, um eine Anleitung zu einem Thema zu erhalten (explanatorische Hilfe: „Wie kann ich telefonieren?“) oder um die Position einer Funktion im grafischen Kontext zu finden (lokalisierende Hilfe: „Wo kann ich die Luftverteilung einstellen?“). Ziel muss es also sein, diese Form der Fragen zu verarbeiten und dementsprechende Hilfeausgaben zu produzieren.

Zu diesem Zweck muss zunächst das Vokabular der Hilfeanfragen näher bestimmt werden. Dazu wurden 7 Versuchspersonen ohne Vorkenntnisse über das iDrive-Bediensystem gebeten, bestimmte Aufgaben mit der Sprachbedienung auszuführen. Dabei wurde darauf geachtet, dass in der Formulierung der Aufgaben keine Sprachkommandos oder Synonyme genannt wurden. Des weiteren erhielten die Benutzer keinerlei Hinweise auf die Hilfefunktionalität des Systems, es wurde lediglich die Aktivierung des SDS mittels der PTT-Taste erläutert. Typisch für die resultierenden Spracheingabeversuche waren folgende Anfragetypen:

1. „Hilfe“ (42% der Anfragen)
2. „Und was muss ich jetzt sagen?“ (25% der Anfragen)
3. „Wie kann ich ...?“ (18% der Anfragen)
4. „Wo ist ...?“ (9% der Anfragen)
5. „Was muss ich tun, um ...?“ (5% der Anfragen)

Während die Äußerung „Hilfe“ zur Ausgabe der allgemeinen Hilfe führt, resultieren die übrigen Anfragen in einer Fehlerkennung, d.h. 58% der Äußerungen beim Erstkontakt führen zu keinem produktiven Dialog. Rechnet man die Äußerung „Und was muss ich jetzt sagen?“ auf Grund der Vagheit der Anfrage der allgemeinen Hilfe zu, kann die Syntax der verbleibenden Anfragen folgendermaßen beschrieben werden:

$$\left[\underbrace{\{\text{Wie}|\text{Wo}|\text{Was}\}}_{\text{Fragewort}} \mid \underbrace{\{\text{kann}|\text{ist}|\text{muss}\}}_{\text{Verb}} \mid \text{Intentionsbeschreibung} \right]$$

Das Vokabular für die Hilfe umfasst die aufgeführten Fragewörter und Verben sowie eine Liste mit möglichen Intentionsbeschreibungen.

Da die natürlichsprachlichen Hilfeanfragen an keinen festen Kontext gebunden sind, muss die Generierung der Hilfeausgaben dynamisch erfolgen. Als Ausgangspunkt für die notwendigen Handlungsanweisungen dient der aktuelle Systemkontext. Die Bestimmung des Zielkontexts erfolgt anhand der Analyse der Intentionsbeschreibung. Diese muss zum einen das reguläre Vokabular des SDS beinhalten, zum anderen jedoch auch Synonyme aus dem umgangssprachlichen Bereich (z.B. „Gebälse“ ergänzend zu „Luftverteilung“). Um eine Handlungsanweisung für die Zielintention ausgehend vom aktuellen Bedienkontext zu erzeugen, wird das SDS als zyklischer, ungerichteter Graph modelliert. Jeder Kontext/Zustand des

SDS wird durch einen Knoten im Graphen $G = (V, E)$ repräsentiert; V entspricht dabei der Knotenmenge, E der Kantenmenge (siehe Abbildung 5.13). Jede Kante besitzt eine Gewichtung w , welche eine positive reelle Zahl sein muss. Aufgabe des

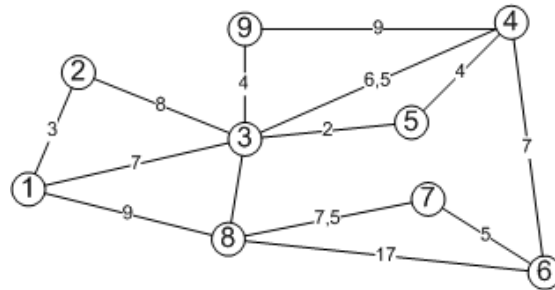


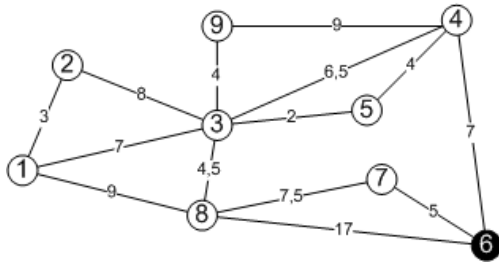
Abbildung 5.13: Zyklischer, ungerichteter Graph: der Graph G besteht aus neun Knoten ($V_1 \dots V_9$) und 14 Kanten ($E_{1,i} \dots E_{j,14}$).

Hilfesystems ist, die kürzeste Verbindung d_{V_S, V_Z} zwischen einem Startknoten V_S und einem Zielknoten V_Z zu ermitteln und so eine optimale Handlungsanweisung zu generieren.

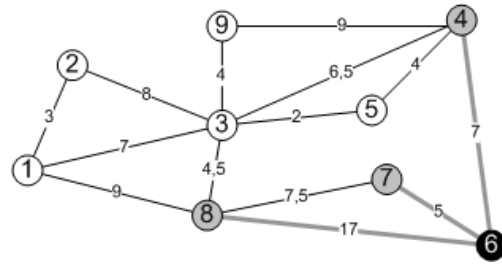
Die kürzeste Verbindung zwischen zwei Knoten in einem zyklischen, ungerichteten Graphen kann über verschiedene Algorithmen ermittelt werden. In Frage kommen z.B. der Algorithmus von Dijkstra oder der A*-Algorithmus. Beide Algorithmen erlauben die Ermittlung des kürzesten Pfades in einem Graphen mit positiven Kantengewichten. Dem A*-Algorithmus liegt eine Heuristik zugrunde, anhand derer eine Schätzung getroffen werden kann, wie hoch die Kosten zur Erreichung des Zielknotens sein werden (vgl. Wikipedia, 2007). Auf diese Weise arbeitet der A*-Algorithmus sehr effizient und laufzeitoptimiert. Für die in dieser Arbeit betrachteten Umfänge spielt die Laufzeit jedoch keine entscheidende Rolle, da die Graphen nur einen Teilbereich des SDS abdecken und somit sehr klein ausfallen. Zur Validierung der Tauglichkeit des Graphenmodells reicht eine einfachere Suchvariante wie der Algorithmus von Dijkstra aus.

Die Funktionsweise dieses Algorithmus ist in Abbildung 5.14 auf Seite 93 veranschaulicht. Sobald das Hilfesystem einen optimalen Pfad berechnet hat, kann ein Dialog mit den entsprechenden Handlungsanweisungen initiiert werden. Zu diesem Zweck besitzt jeder Knoten im Graph verschiedene Informationen:

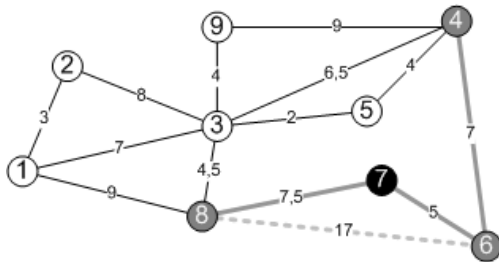
- Feedback (FB): dient zur Validierung der Intention des Benutzers, z.B. „Um jemanden anzurufen“
- Hilfetext (HT): beschreibt eine Handlungsanweisung, um zum aktuellen Knoten zu gelangen, z.B. „nennen Sie einen Namen aus dem Telefonbuch gefolgt von anrufen“
- Beispiel (B): kann optional an den Hilfetext angefügt werden, z.B. „zum Beispiel Markus Abel anrufen“



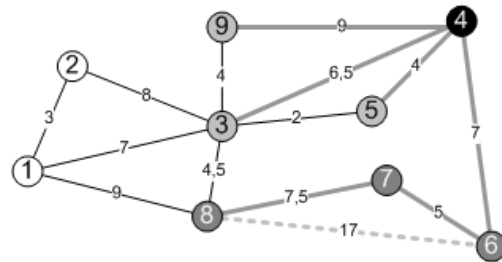
(a) Zu Beginn wird der Graph initialisiert, die Distanzen aller Knoten untereinander werden auf unendlich gesetzt. Startknoten sei V_6 , Zielknoten V_3 .



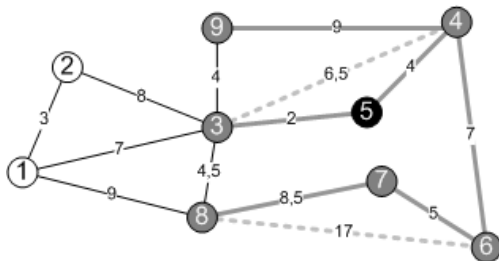
(b) Die jeweils angrenzenden Kanten $E_{6,4}$, $E_{6,7}$ und $E_{6,8}$ werden als kürzeste Entfernung zu den Knoten V_4 , V_7 und V_8 gekennzeichnet.



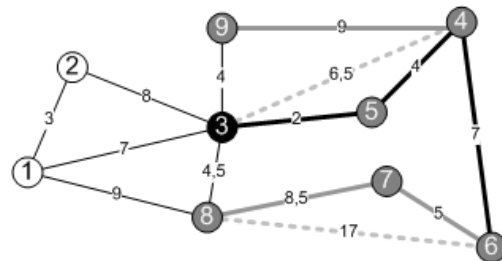
(c) Die kürzeste Verbindung ist $E_{6,7}$, daher wird der angrenzende Knoten V_7 neuer aktueller Knoten. Die angrenzende Kante $E_{7,8}$ stellt die Verbindung $d_{6,7,8}$ her. Da diese kürzer ist als die bereits existierende Verbindung $d_{6,8}$, wird diese als kürzester Pfad ausgeschlossen.



(d) Neue kürzeste Verbindung im betrachteten Graphen ist $d_{6,4}$. V_4 ist neuer aktueller Knoten, die kürzeste angrenzende Verbindung ist $d_{4,5} = 4$.



(e) Da $d_{6,4,5}$ weiterhin die kürzeste Verbindung in G darstellt, wird V_5 neuer aktueller Knoten. Die kürzeste an V_5 angrenzende Kante ist $E_{3,5}$. Da $d_{6,5,3}$ länger ist als die neu gefundene, wird $d_{6,5,3}$ ebenfalls aus der Lösungsmenge ausgeschlossen.



(f) $d_{6,4,5,3}$ ist die neue kürzeste Verbindung im Graphen, d.h. der Zielknoten V_3 wird neuer aktueller Knoten und der Algorithmus wird beendet.

Abbildung 5.14: Algorithmus von Dijkstra zur Bestimmung des kürzesten Pfades zwischen zwei Knoten.

Jede Ausgabe des Hilfesystems besitzt demnach folgende Struktur:

$$| \text{FB} | \underbrace{|\text{HT}(+B)|}_{\text{Anweisungen}} | \text{Spracheingabe} |$$

Die Anweisungen können aus Hilfetexten von verschiedenen Knoten bestehen, wobei nach jeder Anweisung eine Spracheingabe des Benutzers erwartet wird. Abbildung 5.15 verdeutlicht die Generierung einer Anweisung. Der Benutzer befindet sich im Kontext Radio und stellt die Frage „Wie kann ich einen Eintrag aus dem Telefonbuch löschen?“. Das System berechnet den kürzesten Pfad d_{V_1, V_2, V_4} zu V_4 . Zunächst wird das Feedback des Zielknotens V_4 ausgegeben („Um einen Eintrag aus dem Telefonbuch zu löschen“). Anschließend wird der Hilfetext für Knoten V_2 ausgegeben („sagen Sie Telefonbuch“). Der Benutzer muss nun die Spracheingabe „Telefonbuch“ vornehmen. Anschließend fährt das Hilfesystem mit der Anweisung zu Knoten V_4 fort („Nennen Sie einen Eintrag aus dem Telefonbuch gefolgt von löschen“) und fügt das zugehörige Beispiel an („zum Beispiel Markus Abel löschen oder Eintrag 4 löschen“).

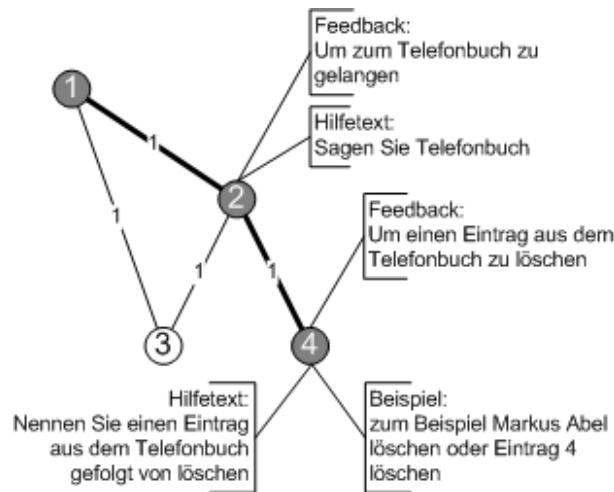


Abbildung 5.15: Anweisung zum Thema „Eintrag aus dem Telefonbuch löschen“: Knoten 1 = Radio, Knoten 2 = Telefonbuch, Knoten 3 = Eintrag anrufen, Knoten 4 = Eintrag löschen

Analog zur kontextspezifischen Hilfe (vgl. Kapitel 5.2.2) wird der Informationsgehalt der Handlungsanweisungen an den Erfahrungsgrad des Anwenders angepasst (siehe Dialog 5.8).

User:	Wie kann ich telefonieren?
System:	〈 <i>Anfänger</i> 〉 Um jemanden anzurufen, nennen Sie einen Namen aus ihrem Telefonbuch gefolgt von anrufen, zum Beispiel Markus Abel anrufen, oder sagen Sie Nummer wählen.
System:	〈 <i>Experte</i> 〉 Nennen Sie einen Namen aus ihrem Telefonbuch gefolgt von anrufen oder sagen Sie Nummer wählen.
User:	Nummer wählen.
System:	〈 <i>Anfänger</i> 〉 Bitte sprechen Sie die Nummer in einzelnen Ziffern.
System:	〈 <i>Experte</i> 〉 Bitte Nummer sprechen.

Dialog 5.8: Anpassung der Anweisung zum Thema „Telefonieren“ für Anfänger und Experten.

5.4 Experimentelle Bestimmung der Systemparameter

In diesem Kapitel werden die Versuche erläutert, die zur Bestimmung der Lern- und Vergessenskurven im Umgang mit SDS im Fahrzeug notwendig sind (vgl. Kapitel 5.2.1.1 und 5.2.1.2). Basierend auf den Ergebnissen kann entschieden werden, welche der mathematischen Modelle bezüglich Lernen und Vergessen von Sprachkommandos am besten für den Einsatz im Fahrzeug geeignet sind.

5.4.1 Lernen vom Sprachkommandos

5.4.1.1 Testdesign

Die Tests fanden in einem BMW 6er mit iDrive-Bediensystem statt. Als Teststrecke wurde ein Rundkurs auf dem BMW Werksgelände in Regensburg gewählt. Das Werksgelände ist gut geeignet, um verschiedene anspruchsvollere Verkehrssituationen in einer kontrollierten Umgebung nachzustellen, z.B. Tempo 30 Zonen, rechts vor links Regelungen oder verengte Fahrbahnen mit Hindernissen.

Die Testgruppe bestand aus sieben Versuchspersonen (VP) im Alter zwischen 26 und 43 Jahren. Die Auswahl der VP erfolgte nach den Kriterien Erfahrung mit SDS und Fahrerfahrung. Für die Bestimmung des Lernverhaltens beim Erstkontakt mit einem SDS dürfen die VP keine bzw. sehr wenig Erfahrungen diesbezüglich gesammelt haben. Um Interferenzen zwischen der Fahrzeugführung und der Systembedienung zu vermeiden, mussten die Probanden eine Fahrleistung von bisher 100.000 km oder alternativ fünf Jahre Fahrerfahrung aufweisen. Zudem musste die jährliche Fahrleistung mindestens 10.000 km betragen.

Vor Beginn der Tests erhielten die VP eine Einführung in die Fahrzeugbedienung, anschließend wurde eine kurze Eingewöhnungsfahrt durchgeführt. Während des darauffolgenden Tests bestand die Aufgabe der VP darin, zehn Aufgaben mit

unterschiedlichem Schwierigkeitsgrad zu bearbeiten (siehe Tabelle 5.3). Die Komplexität der Aufgaben wurde anhand der zur Lösung minimal notwendigen Dialogschritte gemessen. Die VP erhielten jeweils eine Beschreibung der zu bearbei-

Aufgabe	Komplexitätsgrad
Frequenz 103,0	3
Zielliste	3
Zieleingabe	3
Zielführung starten	3
Navigationshinweise aus	4
Perspektive	4
Maßstab 1km	4
Autobahn vermeiden	4
CD ein	3
Bordcomputer	3

Tabelle 5.3: Aufgaben der Lerntests mit zugehörigen Komplexitätsgraden

tenden Aufgabe und mussten das entsprechende Sprachkommando in der Hilfe suchen. Nach der Lösung der Aufgabe wurde die nächste Aufgabe gestellt. Die Aufgaben wurden solange wiederholt, bis die zugehörigen Sprachkommandos memorisiert wurden. Für jeden Durchlauf wurde die Zeitspanne zwischen der Aktivierung der Spracheingabe und der Äußerung des korrekten Sprachkommandos gemessen. Die ermittelten Zeitspannen wurden anhand des zuvor festgelegten Komplexitätsgrades standardisiert.

5.4.1.2 Ergebnisse

Generell spiegeln die Ergebnisse die Erwartung wieder, dass der Lernerfolg anfangs groß ist, d.h. das Lernen erfolgt zunächst sehr schnell. Mit zunehmenden Wiederholungen nimmt die zur Lösung einer Aufgabe notwendige Zeit stetig ab und nähert sich einer horizontalen Asymptote (vgl. Abbildung 5.16). Ob eine Potenzfunktion oder eine Exponentialfunktion den Lernprozess besser beschreibt (vgl. Kapitel 5.2.1.1), kann mit Hilfe des χ^2 -Goodness-of-Fit-Tests ermittelt werden (Rasch et. al., 2004). Je mehr χ^2 gegen Null tendiert, desto weniger weichen die geschätzten Werte f_e von den beobachteten Werten f_o ab.

$$\chi^2 = \sum_{i=1}^k \frac{(f_o - f_e)^2}{f_e} \quad (5.11)$$

Die χ^2 -Werte für die Potenzfunktion und die Exponentialfunktion nehmen den jeweils niedrigsten Wert mit den in Tabelle 5.4 Lernkoeffizienten an. Das bedeutet,

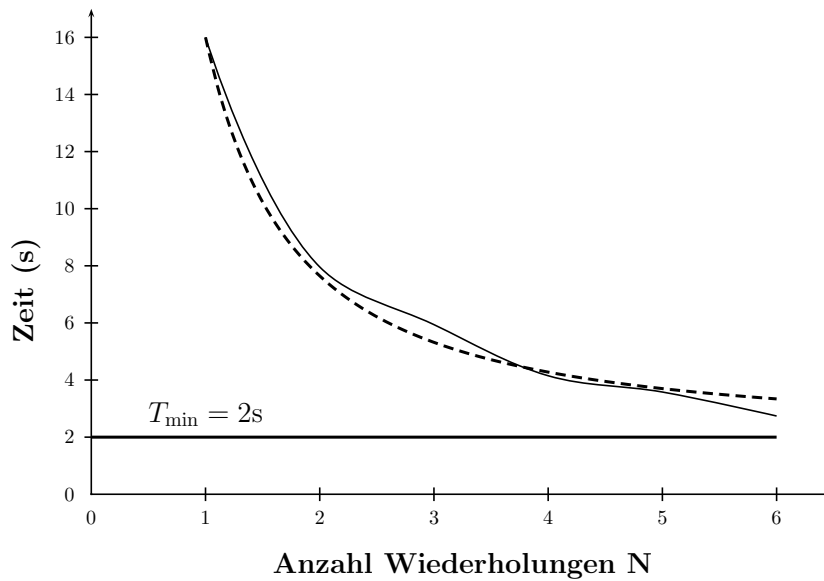


Abbildung 5.16: Resultierende Lernkurve und Lernkurve basierend auf einer Potenzfunktion (gestrichelte Linie).

Funktion	χ^2	Lernkoeffizient α
Exponentialfunktion	2,72	0,41
Potenzfunktion	1,31	0,42

Tabelle 5.4: χ^2 -Werte der Lernfunktionen

dass die Potenzfunktion (vgl. Gleichung 5.1) mit einem Lernkoeffizienten von $\alpha = 1,31$ die beobachteten Ergebnisse besser beschreibt als die Exponentialfunktion, da χ^2_{min} für die Potenzfunktion kleiner ist als für die Exponentialfunktion.

5.4.2 Ausbildung von Transferwissen

Die Versuche zur Ausbildung von Transferwissen fanden in einem BMW 3er statt. Die Versuche fanden im stehenden Fahrzeug statt. Die Testgruppe bestand aus sieben VP im Alter zwischen 20 und 24 Jahren. Alle VP hatten keinerlei Erfahrung im Umgang mit SDS.

Ziel der Tests war es, zu ermitteln, wie stark Wissen über bestimmte Funktionen und Konzepte des SDS auf ähnliche Bereiche übertragen werden kann. Dazu erhielten die Probanden eine Liste mit zehn Sprachkommandos, welche sie zweimal sprechen mussten, um ein Gefühl für die Funktionsweise des SDS zu bekommen. Anschließend erhielten die Probanden Aufgaben, bei denen sie ähnliche Funktio-

nen anwenden mussten (siehe Tabelle 5.5). Die entsprechenden Sprachkommandos mussten von den angewendeten Konzepten abgeleitet werden.

Angewendete Funktionen	Ähnliche Funktionen	Kommandos
CD ein CD Titel 5	CD deaktivieren Beliebiger anderer CD Titel	CD aus CD Titel 9
Nächster Titel Vorherige CD	Titel davor abspielen nachfolgende CD abspielen	Vorheriger Titel Nächste CD
Verkehrsfunk aus Mautstrecken vermeiden aus Zufallswiedergabe ein	Verkehrsfunk aktivieren Mautstrecken aktivieren	Mautstrecken vermeiden ein Zufallsweidergabe aus
Autobahn vermeiden	Random-Wiedegabe deaktivieren Autobahn verwenden	Autobahn vermeiden aus
Bildschirm aus	Display aktivieren	Bildschirm ein

Tabelle 5.5: Aufgaben der Versuche zum Transferwissens

Als Maß für die Transferleistung wurde die Anzahl an korrekt abgeleiteten Kommandos herangezogen. Die Auswertung der Test ergab, dass die Probanden in 89% der Fälle die gesuchten Kommandos von den bereits bekannten Funktionen und Konzepten ableiten konnten. Daraus kann geschlossen werden, dass vorhandenes Wissen über Funktionen und Konzepte einen wichtigen Beitrag zum Erlernen weiterer Funktionalitäten leistet.

5.4.3 Vergessen von Sprachkommandos

Die Untersuchung zum Vergessen von Sprachkommandos fand zum Teil in einem BMW 3er und zum Teil an einem Laptop statt. Die Versuche wurden im stehenden Fahrzeug durchgeführt, da auf Grund der angewendeten Testmethodik keine mentale Belastung zur Messung der Vergessensleistung nötig ist.

Die Testgruppe bestand aus zehn VP im Alter zwischen 23 und 53 Jahren. Die Auswahl der VP erfolgte nach dem Kriterium der Erfahrung im Umgang mit SDS im Allgemeinen und dem iDrive SDS im Besonderen. Für die Memorierungstests ist es von entscheidender Bedeutung, dass die VP keine Sprachkommandos für iDrive kennen.

Die Memorierungstests bestanden aus zwei Phasen, einer Lernphase und einer Vergessensphase. In der Lernphase erlernten zehn VP zwei Blöcke mit jeweils

10 Sprachkommandos durch Anwendung im Fahrzeug (siehe Tabelle 5.6). Die

Funktionsblock 1		Funktionsblock 2	
Task 1	CD Player	Task 11	TV einschaltenn
Task 2	Track 5 von CD hören	Task 12	TV Sender 'ARD' ansehen
Task 3	Radio hören	Task 13	Luftverteilung ändern
Task 4	Radiostation 'Antenne Bayern' hören	Task 14	Uhr stellen
Task 5	Frequenz 103,0 MHz hören	Task 15	Datum stellen
Task 6	Klangeinstellungen	Task 16	Bildschirmhelligkeit einstellen
Task 7	Zielführung an	Task 17	BMW Online
Task 8	Kartenmaßstab 1 km	Task 18	Telefonmenü
Task 9	Routenkriterium 'Autobahn vermeiden'	Task 19	Assistenzfenster
Task 10	Routenkriterium 'Fahren vermeiden'	Task 20	Bildschirm ausschalten

Tabelle 5.6: Aufgaben für die Memorierungstests

entsprechenden Sprachbefehle wurden den Probanden in Form einer Liste zur Verfügung gestellt. Die Sprachkommandos wurden so lange eingeübt, bis sie frei reproduziert werden konnten. Diese 20 Kommandos bildeten die Basis der weiteren Memorierungstests.

Ziel war es, die Geschwindigkeit des Vergessens zu ermitteln. Zu diesem Zweck wurden über einen Zeitraum von 50 Tagen mehrere Vergessenstests in einer Laborumgebung durchgeführt. In den Tests sollte die Situation nachgestellt werden, dass ein Benutzer eine bestimmte Aktion (z.B. Radio hören) mit dem SDS ausführen will. Um jeglichen Einfluß auf das verbale Gedächtnis auszuschließen, wurden die Intentionen für den Benutzer nicht vorgelesen, sondern in Form von Piktogrammen dargestellt (siehe Abbildung 5.17). Jedes Piktogramm soll eine Inten-



Abbildung 5.17: Piktogramme zur Bildung der Intentionen 'Telefon', 'Autobahn vermeiden' und 'Radio'. Quelle: <http://images.google.de>

tion beim Benutzer erzeugen, die dann einer Aufgabe zugeordnet wird und den Abruf des entsprechenden Sprachkommandos aus dem Gedächtnis zur Folge hat. Die Memorierungstests wurden nach neun, 22, 36 und 50 Tagen durchgeführt. Als

Maß für das Vergessen wurden die korrekt wiedergegebenen Sprachkommandos verwendet.

Die resultierende Vergessenskurve ist in Abbildung 5.18 zu sehen. Erwartungsgemäß

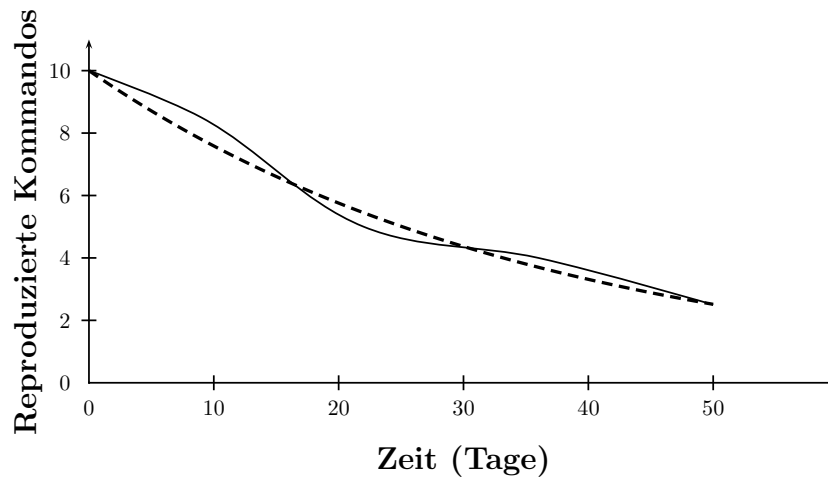


Abbildung 5.18: Resultierende Vergessenskurve und Vergessenskurve basierend auf einer Exponentialfunktion (gestrichelte Linie).

ist die Zahl der erinnerten Kommandos anfänglich relativ hoch und nimmt mit zunehmender Zeit stetig ab. Die Resultate zeigen, dass nach 50 Tagen ca. 75% der Sprachkommandos vergessen wurden. Um zu bestimmen, welche der in Kapitel 5.2.1.2 erläuterten mathematischen Funktionen den Vergessensprozeß im Fahrzeug bestmöglich beschreibt, wurde ebenfalls der χ^2 -Goodness-of-Fit-Test herangezogen. Die ermittelten minimalen χ^2 -Werte sind in Tabelle 5.7 aufgeführt. Den

Funktion	χ^2	Vergessenskoeffizient δ
Logarithmische Funktion	2,11	0,58
Exponentialfunktion	0,12	0,027
Potenzfunktion	1,77	0,22
Exponentialfunktion mit Wurzel	0,98	0,15

Tabelle 5.7: χ^2 -Werte der Vergessensfunktionen

besten χ^2 -Wert liefert die Exponentialfunktion (vgl. Gleichung 5.7), d.h. diese weist die beste Anpassungsgüte auf.

5.5 Auswahl der Beschreibungsmodelle

Nach Auswertung der Versuche bezüglich des Lernens und Vergessens von Sprachkommandos, kann die Frage nach den am besten geeigneten mathematischen Be-

schreibungsmodellen aus Kapitel 5.2.1.1 und 5.2.1.2 beantwortet werden. Der Lernprozess kann am besten mit Hilfe des Potenzgesetzes der Übung beschrieben werden. Um zu bestimmen, nach wie vielen Wiederholungen N ein Sprachkommando erlernt wurde, kann demnach folgende Formel verwendet werden:

$$N = \sqrt[{-1,31}]{\frac{T}{B}} \quad (5.12)$$

Überträgt man diese Gleichung in ein Diagramm, kann daraus für jede Zeit bei der ersten Bedienung einer Funktion der zugehörige Schwellenwert abgelesen werden. Abbildung 5.19 verdeutlicht, dass bei einer erstmaligen Bediendauer von 30 Sekunden der zugehörige Schwellenwert $N = 7,9$ beträgt.

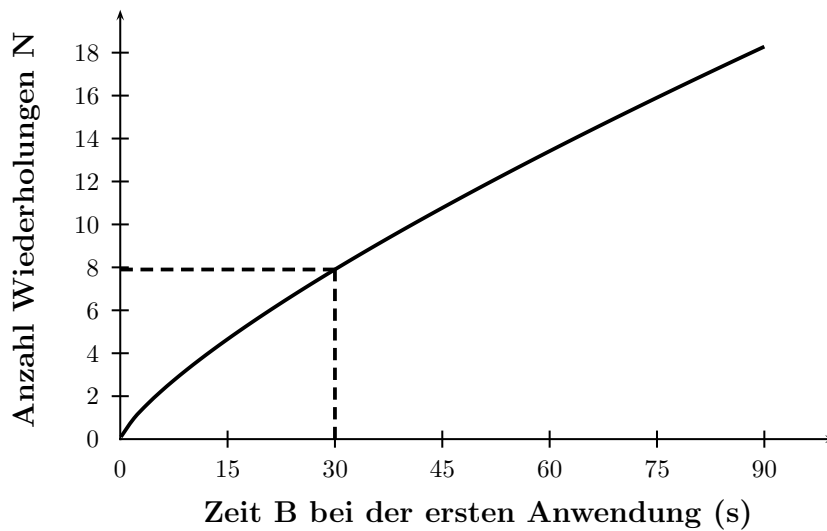


Abbildung 5.19: Berechnung des Schwellenwerts N

Durch jede direkte Verwendung eines Sprachbefehls erhöht sich der entsprechende Index i um den Wert 1. Ist dieser Befehl Bestandteil einer Funktionsfamilie ϕ , so werden die Indices aller darin enthaltenen Befehle basierend auf den Ergebnissen zur Anwendung von Transferwissen (siehe Kapitel 5.4.2) um den Wert $\gamma = 0,89$ erhöht. Erreicht der Index einer Funktion den zugehörigen Schwellenwert N , so gilt diese Funktion als gelernt.

Das Vergessen kann mit Hilfe einer Exponentialfunktion beschrieben werden:

$$i_{neu} = i_{alt} \cdot e^{-0,027 \cdot T} \quad (5.13)$$

Ausgehend vom letzten Tag der Verwendung einer Funktion nimmt der Index entsprechend der Exponentialfunktion solange ab, bis diese erneut verwendet wird.

Die Veränderung der Indices unter Berücksichtigung aller Beschreibungsmodelle ist beispielhaft in Abbildung 5.20 dargestellt.

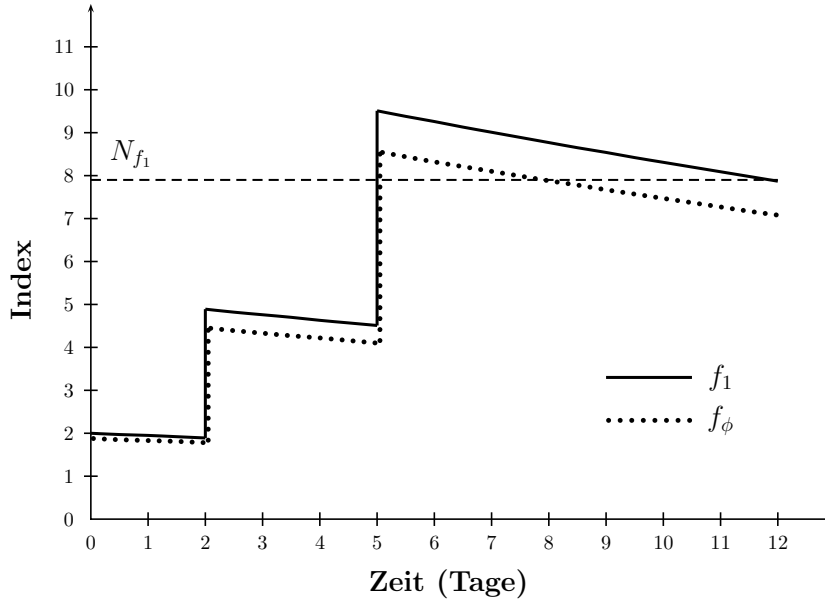


Abbildung 5.20: Verlauf der Indices einer Funktion f_1 sowie der Funktionen f_ϕ einer Funktionsfamilie ϕ : Zunächst wird die Funktion f_1 zweimal verwendet, so dass $i_{f_1} = 2$. Gleichzeitig erhöhen sich die Indices der restlichen Funktionsfamilie, so dass $i_{f_\phi} = 2 \cdot 0,89 = 1,88$. Während der folgenden zwei Tage wird f_1 nicht verwendet, d.h. alle Indices nehmen langsam wieder ab. Erst bei der nächsten Verwendung (dreimal an Tag 2) erhöhen sich die Indices der Funktionen wieder, so dass $i_{f_1} = 4,89$ und $i_\phi = 4,45$. Sobald der Schwellenwert N_{f_1} erreicht wird, gilt die Funktion f_1 als gelernt und wird in der kontextspezifischen Hilfe ans Ende der Hilfeausgabe verschoben.

5.6 Fazit

Die kontextspezifische Hilfe ist durch eine Reduzierung der angebotenen Informationseinheiten gekennzeichnet. Pro Dialogschritt werden maximal vier (für Anfänger) oder fünf (für Experten) Sprachbefehle ausgegeben, daraus resultiert eine Aufteilung eines Hilfedialogs in mehrere Ebenen. Die Einstufung eines Benutzers als Anfänger oder Experte erfolgt durch ein Benutzermodell, welches den Erfahrungsgrad des Anwenders anhand der Kriterien Spracherkennungsfehler, Hilfe- und Optionsaufrufe, Häufigkeit von Barge-In und Timeouts bestimmt. Die Inhalte der Hilfe werden an das Wissen des Benutzers angepasst, indem das Lernen und Vergessen von Sprachkommandos mit Hilfe des Potenzgesetzes der Übung und einer Vergessenskurve modelliert wird. Der Fokus der Hilfe liegt auf Kommandos, die der Benutzer weniger häufig verwendet oder gar nicht kennt. Hilfe zu Funktio-

nen, die häufig verwendet werden und gut bekannt sind, werden an das Ende der Hilfeausgaben verschoben. Die kontextspezifische Hilfe beinhaltet zudem eine visuelle Komponente, die dem Benutzer haptische Eingaben erlaubt. Die Auswahl eines Sprachkommandos kann während der Hilfeausgaben per Sprache (Barge-In) oder Controller (Push-In) erfolgen. Nach einer haptischen Interaktion wird der Sprachdialog regulär fortgesetzt.

Die kontextunabhängige Hilfe bietet dem unerfahrenen Benutzer Hilfestellung in unbekannten Domänen des SDS. Die Hilfeanfrage kann als Frage formuliert werden (z.B. „Wie kann ich Radio hören?“), worauf das Hilfesystem eine entsprechende Anleitung ausgibt. Dabei wird der optimale Pfad zwischen dem aktuellen Dialogzustand und dem Zielstand mit Hilfe des Algorithmus von Dijkstra ermittelt.

Kapitel 6

Implementierung des Prototyps

In diesem Kapitel werden die technischen Aspekte des Hilfesystems näher beleuchtet. Details hinsichtlich der Programmierung werden nicht behandelt, da die Implementierung in Zusammenarbeit mit einem externen Dienstleister erfolgte und keinen Hauptbestandteil dieser Arbeit darstellt. Die Ausführungen beschränken sich auf die Erläuterung der Schlüsselkonzepte.

Die Umsetzung der Konzepte für ein adaptives Hilfesystem erfolgte auf Basis einer bestehenden Simulation für den iDrive Nachfolger. Somit ist sichergestellt, dass sich das Hilfesystem nahtlos in das bestehende Anzeige-Bedienkonzept integriert und damit mehr als nur ein rudimentärer Prototyp ist. Zudem bietet die vorhandene Simulation alle notwendigen Schnittstellen und Komponenten, die für das Hilfesystem notwendig sind. Nach einer kurzen Erläuterung der Systemarchitektur in Kapitel 6.1, werden die entscheidenden Komponenten zur Steuerung der kontextspezifischen Hilfe in Kapitel 6.2 präsentiert. Die Beschreibung der Implementierung der kontextunabhängigen Hilfe erfolgt in Kapitel 6.3.

6.1 Systemarchitektur der Dialogsimulation

Die Spracherkennungskomponente der Simulation besteht aus einem TEMIC-Spracherkenner der aktuellsten Generation. Der Erkenner ist semiadaptiv, d.h. dass er sich mit der Zeit an individuelle Sprechereigenschaften anpasst. Dies resultiert in einer anfänglich etwas schlechteren Erkennungsleistung, mit zunehmender Verwendung steigt diese jedoch an und übertrifft bisherige TEMIC-Spracherkenner hinsichtlich ihrer Erkennrate.

Der TEMIC-Spracherkenner ist grammatikbasiert und verarbeitet Java Speech Grammar Format (JSGF) Grammatiken (Sun Developer Network, 1998). Das Vokabular besteht aus ca. 5000 Begriffen. Die semantische Verarbeitung der Eingaben, die Simulation verschiedener Geräte sowie die Dialogsteuerung erfolgen über ein in C++ implementiertes Dialogsystem. Die Planung der Sprachausgaben erfolgt mit Hilfe von Excel-Dateien im CSV-Format. Diese Dateien enthalten die

Texte, welche durch die TTS-Engine RealSpeak Solo von Nuance synthetisiert werden.

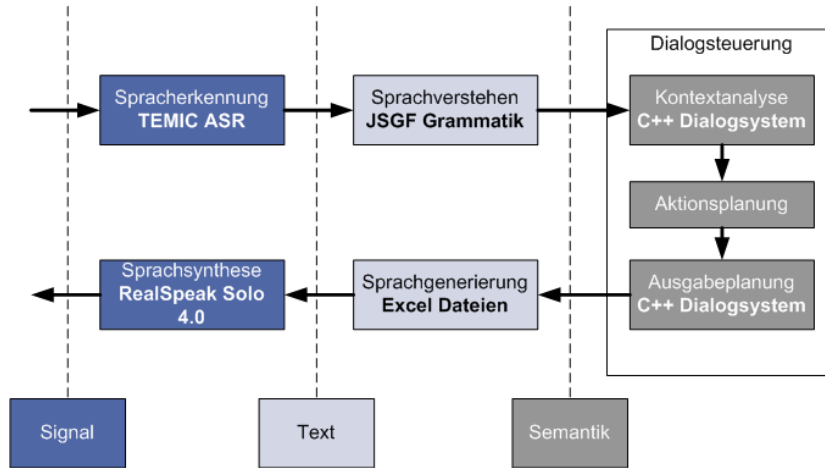


Abbildung 6.1: Systemarchitektur des Sprachdialogsystems in Anlehnung an Keller (2004) und McTear (2002)

Das visuell-haptische Bediensystem besteht aus einem Controller, mit dessen Hilfe Eingaben durch Drehen und Drücken vorgenommen werden können. Jede Betätigung löst ein Ereignis aus, welches im C++ Dialogsystem verarbeitet wird und letztendlich eine Veränderung der graphischen Anzeige bewirkt. Die Simulation der Bedienoberfläche erfolgt mittels Macromedia Director.

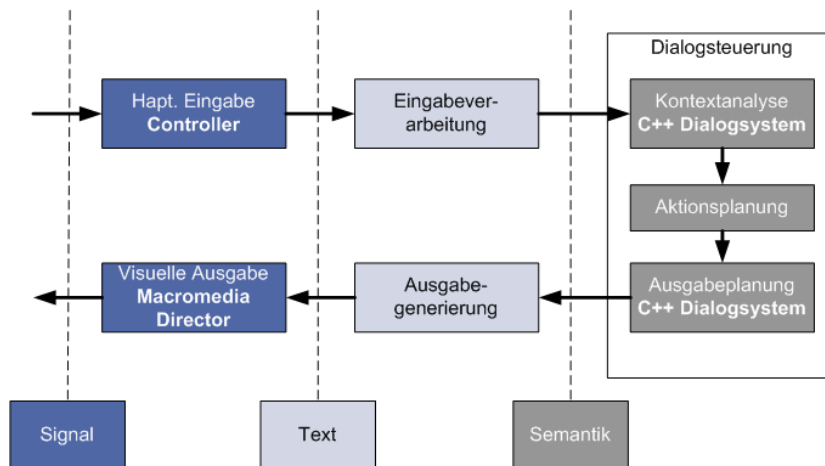


Abbildung 6.2: Systemarchitektur des visuell-haptischen Bediensystems

6.2 Kontextabhängige Hilfe

Die Implementierung des in Kapitel 5.5 erläuterten Adaptionismus erfordert eine Anpassung der Dialogsteuerung. Zunächst ist es notwendig, verschiedene Parameter während des Umgangs mit dem SDS zu bestimmen.

Für die Generierung des Benutzermodells muss zum einen die Zeit für die erste Benutzung einer Funktion B ermittelt werden, um berechnen zu können, nach wie vielen Wiederholungen diese Funktion als gelernt betrachtet werden kann (vgl. Kapitel 5.2.1.1). Die Messung dieser Zeitspanne stellt eine große Herausforderung dar, da zunächst definiert werden muss, mit welchem Sprachkommando eine bestimmte Aktion begonnen wird. Nicht immer ist das Verhalten eines Benutzers zielgerichtet. Exploratives Verhalten kann die Zeitmessung unter Umständen verfälschen. Für die Bestimmung von B müssen Aktionssequenzen hinterlegt werden, welche die möglichen Pfade zur Ausführung einer Funktion beschreiben. Eine starre Definition von Aktionssequenzen ist jedoch sehr aufwändig und nicht zielführend, da jede mögliche Aktionssequenz einzeln aufgeführt werden muss. So kann die Funktion „Maßstab 500m“ auf verschiedene Arten aufgerufen werden, z.B.:

```
Navigation → Karte → Optionen → Maßstab 500m
Navigation → Optionen → Karte → Optionen → Maßstab 500m
Karte → Maßstab 500m
```

Theoretisch können nach jedem Kommando die Optionen aufgerufen werden; es besteht zudem die Möglichkeit, dass gerade Anfänger die Optionen mehrmals hintereinander verwenden. Jede mögliche Kombination von Sprachkommandos bezogen auf ein Dialogziel müsste einzeln in einer Aktionsbibliothek hinterlegt werden. Legt man den Berechnungen 700 bis 1000 sprachbedienbare Funktionen zu Grunde, kann unter der Annahme, dass jede Funktion ca. 5-10 mögliche Pfade besitzt, davon ausgegangen werden, dass eine Aktionsbibliothek aus ca. 3500 bis 10000 Aktionssequenzen bestünde. Um eine Verarbeitung in Echtzeit garantieren zu können, müsste die Aktionsbibliothek in einer Datenbank gespeichert werden. Diese verbraucht jedoch zusätzliche System- und Speicherressourcen, was für eine Verwendung im Fahrzeug nachteilig ist.

Eine wesentlich effizientere Methode der Zeitermittlung kann durch die Bildung eines Intentionspools erfolgen. Dabei wird für jeden Endzustand, der im SDS erreicht werden kann, der längstmögliche Pfad hinterlegt. Jede sprachliche Interaktion des Benutzers wird auf einem History-Stack abgelegt, welcher die zuletzt aufgerufenen Sprachkommandos beinhaltet. Durch einen Abgleich des History-Stacks mit dem Intentionspool können die Zeiten B für verschiedene Endzustände berechnet werden (siehe Abbildung 6.3). Folgen zwei mögliche Endzustände aufeinander (z.B. „Karte“ und „Karte nordweisend“), stellt sich die Frage, inwiefern das nachfolgende Kommando noch in Relation zum vorangegangenen steht. Im Rah-

men dieser Arbeit wird dabei auf Basis einer Schätzung die Annahme getroffen, dass ein Kommando, das innerhalb von 15 Sekunden oder weniger nach einem Endzustand gesprochen wird, noch der Aktionssequenz hinzugerechnet wird.

Die Reihenfolge der verwendeten Sprachkommandos ist bei dieser Methode der Zeiterfassung von untergeordneter Bedeutung. Eine im Intensionspool hinterlegte Sequenz kann auch von Funktionen unterbrochen werden, die nicht im Pfad enthalten sind. Somit wird einem eventuellen explorativen Verhalten Rechnung getragen. Der Aufruf der Hilfe bzw. der Optionen wird nicht explizit im Intensionspool hinterlegt. Generell gilt, dass die bei deren Aufruf resultierenden Zeiten stets mit in die Zeitmessung einfließen, da auch die Verwendung der Hilfsfunktionen Teil des Lernprozesses ist.

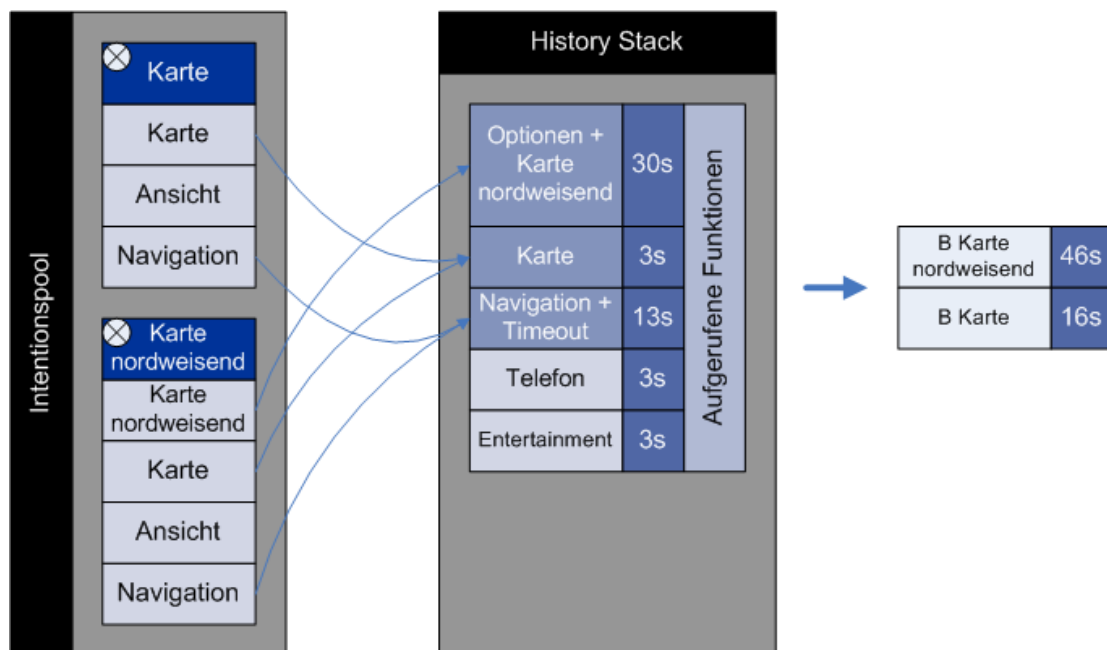


Abbildung 6.3: Ermittlung der Zeit bei der ersten Verwendung einer Funktion: Im Intensionspool werden für jeden möglichen Endzustand die längstmöglichen Pfade hinterlegt. Während der Verwendung des SDS werden die aufgerufenen Funktionen auf einem History-Stack abgelegt. Durch einen Abgleich des History-Stacks mit dem Intensionspool lassen sich die Zeiten B ermitteln. Timeouts und Optionen werden bei der Zeitmessung stets hinzugerechnet.

Neben der Zeit B müssen weitere Parameter ermittelt werden. Für die Bestimmung, ob eine Funktion gelernt wurde, muss die Anzahl der Verwendungen für jede Funktion ermittelt und in einem Index i gespeichert werden. Daneben ist es für die Integration des Modells über das Vergessen wichtig, das Datum der letzten Verwendung einer Funktion zu hinterlegen (vgl. Kapitel 5.2.1.2).

Der Berechnung des generischen Benutzermodells liegen in Anlehnung an Hassel (2006) folgende Parameter zu Grunde:

- Hilfeanfragen
- Optionsanfragen
- Timeouts
- Erkennenfehler
- Anzahl Barge-In

Diese Parameter werden für die Funktionsbereiche Telefon, Entertainment, Navigation und Sonstige erhoben. Aus diesen Informationen kann ein Wert für das Benutzermodell errechnet werden, welcher Rückschlüsse auf den Erfahrungsgrad eines Benutzers zulässt.

Alle Parameter, deren permanente Speicherung im SDS notwendig ist, werden in einer XML-Datei hinterlegt. Diese Datei enthält zu Beginn Informationen über den Erfahrungsgrad eines Benutzers (0=Anfänger, 1=Experte). Anschließend folgen die einzelnen Funktionen des SDS, wobei jeder Eintrag aus einem Index i , dem Datum der letzten Verwendung $time$ und dem Schwellenwert N besteht (siehe Quelltext 6.1).

```
<userstatus>
  <phone>1</phone>
  <nav>0</nav>
  <enter>1</enter>
  <others>0</others>
</userstatus>
<intention0>
  <index>5.80</index>
  <time>20061110</time>
  <N>3.00</N>
</intention0>
<intention1>
  <index>7.46</index>
  <time>20061110</time>
  <N>5.34</N>
</intention1>
<intention2>
  <index>1.41</index>
  <time>20061106</time>
  <N>7.56</N>
</intention2>
```

Quelltext 6.1: User Model XML

Die statischen Parameter des Hilfekonzpts (Gewichtsvektor \overrightarrow{UM}_G , Lern- und Vergessenskoeffizient, etc.) werden in einer ini-datei gespeichert. Auf diese Wei-

se können Korrekturen auf einfache Art und Weise vorgenommen werden (siehe Quelltext 6.2).

```
[ Benutzer ]
    ID=1

[ Generisches Usermodel ]
[ Schwellenwert ]
    wert=0.4
[ Unterschreitungen ]
    count=3
[ Gewicht Hilfsanfragen ]
    gh=0.23
[ Gewicht Optionsanfragen ]
    go=0.30
[ Gewicht Timeouts ]
    gt=0.11
[ Gewicht ASR Fehler ]
    ga=0.62
[ Gewicht BargeIn ]
    gb=-0.37

[ Knowledge Agent ]
[ Memory Factor ]
    alpha=1.31
[ BestTaskTime ]
    sekunden=2
[ Dialog Pause ]
    sekunden=15
[ Vergesskoeffizient ]
    beta=0.027
[ Indexerhöhung ähnliche Intentionen ]
    y=0.89
```

Quelltext 6.2: Statische Systemparameter

6.3 Kontextunabhängige Hilfe

Die Umsetzung des Konzepts der kontextunabhängigen Hilfe erfolgte im Gegensatz zur kontextunabhängigen Hilfe nicht vollständig. In Anlehnung an Nielsen (1994) stellt die kontextunabhängige Hilfe lediglich ein Szenario dar (siehe Abbildung 6.4). Für das Szenario wurden einfache Use Cases ausgewählt, die über ein Shortcut aufgerufen werden können:

- Teilnehmer anrufen

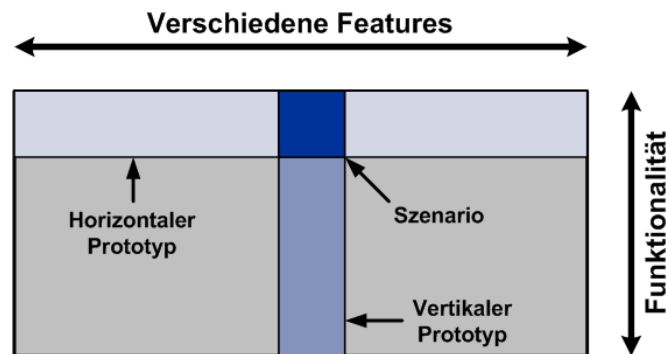


Abbildung 6.4: Klassifikation von Prototypen nach Nielsen (1994)

- Radio hören
- Musik hören
- Klangeinstellungen vornehmen
- Routenkriterien verändern
- Kartendarstellung verändern

Somit entfallen anspruchsvolle Berechnungen hinsichtlich des kürzesten Pfades zu einer Zielfunktion. Da primär untersucht werden soll, inwiefern eine sprachliche Anleitung dazu dienen kann, den Einstieg in das SDS zu erleichtern, spielt die Komplexität der ausgewählten Use Cases für nachfolgende Untersuchungen nur eine untergeordnete Rolle.

Da für die Verarbeitung natürlichsprachlicher Eingaben kein entsprechender SLM-Spracherkenner zur Verfügung stand, wurde die Grammatik des TEMIC-Erkenners für die Verarbeitung natürlichsprachlicher Eingaben angepasst. Dazu werden überflüssige Eingaben und Füllwörter mit Hilfe eines Garbage-Filters entfernt. Der Aufruf eines Hilfetextes erfolgt nach dem in Kapitel 5.3 beschriebenen Muster in Form einer Frage. Die möglichen Kommandos werden in der Grammatik einzeln hinterlegt:

```
<GARBAGE>: (hallo|ich|will|möchte|würde|hätte|gerne|bitte|danke
|jetzt|sofort|plötzlich|gleich|schnell|aber|mal);
```

```
<INTRODUCTION_>: !optional(Wie kann ich)
<INTRODUCTION__> {INT=#2.INT; PAR=#2.PAR;};
```

```
<INTRO_MAP>:
((mir eine|eine|zur|die|andere)
(Karte|Kartenansicht|Kartendarstellung)
(anzeigen lassen|anzeigen|wechseln|ändern))
```

```
{PAR="NO_PARAM"; INT="INTRO_MAP";};
```

Die Hilfetexte für Anfänger und Experten werden in einer Excel-Datei hinterlegt und von der RealSpeak Solo TTS-Engine synthetisiert.

6.4 Fazit

Die größte Herausforderung im Rahmen der Implementierung besteht darin, die notwendigen Parameter wie z.B. die Zeit für die erste Verwendung B zu ermitteln. Herkömmliche Methoden der Planerkennung können aufgrund der Komplexität des SDS nicht angewendet werden. Die Ermittlung der Zeit B geschieht durch die Definition des längstmöglichen Pfades für jeden Endzustand im SDS in einem Intensionspool. Während der Verwendung des SDS werden die aufgerufenen Funktionen auf einem History-Stack abgelegt und mit dem Intensionspool abgeglichen. Die Zeiten für auftretende Timeouts und Optionsaufrufe werden der Zeit B hinzugerechnet.

Diejenigen Parameter, die permanent gespeichert werden müssen, werden in einer XML-Datei abgelegt. Dazu gehören neben B auch die Indices und das Datum der letzten Verwendung der einzelnen Funktionen des SDS. Aus diesen Daten berechnet das Hilfesystem den Schwellenwert N , nach wie vielen Anwendungen eine Funktion als gelernt betrachtet werden kann. Auch dieser Wert wird zusammen mit dem Erfahrungsgrad des Benutzers in verschiedenen Systembereichen (Navigation, Entertainment, Telefon, Sonstige) in der XML-Datei abgespeichert. Die Konfiguration des Hilfesystems kann mit Hilfe einer ini-Datei durchgeführt werden. Darin sind u.a. die Parameter des Gewichtsvektors \overrightarrow{UM}_G sowie der Lern- und der Vergessenskoeffizient hinterlegt.

Das kontextunabhängige Hilfesystem wurde anhand einfacher Use Cases implementiert, aufwändige Berechnungen entfallen somit beim Prototyp. Die Grammatik des TEMIC-Spracherkenners wurde durch Hinzufügen eines Garabage-Filters und der Erweiterung des Vokabulars derart angepasst, dass natürlichsprachliche Systemanfragen möglich sind.

Kapitel 7

Evaluierung des Hilfesystems

Der Evaluierung der in dieser Arbeit vorgestellten Hilfesysteme liegen verschiedene Arbeitshypothesen zugrunde, welche in Kapitel 7.1 erläutert werden. Im Anschluß wird die Evaluierung der Hilfesysteme ausgeführt, welche im Rahmen der Masterarbeit von Stefan Sagstetter erfolgte. Für zusätzliche Informationen zur Evaluierung der Hilfesysteme sei an dieser Stelle auf Sagstetter (2007) verwiesen.

Kapitel 7.2 behandelt die Wahl des Testdesigns, der genaue Ablauf der Versuche ist in Kapitel 7.3 beschrieben. Die Auswertung der Versuche erfolgt in Kapitel 7.4 und umfasst eine Bewertung der Systeminteraktionen, eine Analyse des Blickverhaltens, die subjektive Bewertung der getesteten Systeme, eine Bewertung der Usability anhand des PARADISE Evaluierungsframeworks sowie die Messung der mentalen Belastung während der Aufgaben. Abschließend werden die Ergebnisse in Kapitel 7.5 mit Bezug auf die aufgestellten Arbeitshypothesen diskutiert.

7.1 Arbeitshypothesen

1. Das kontextspezifische Hilfesystem basierend auf adaptiven Komponenten und visueller Unterstützung führt im Gegensatz zum iDrive Hilfesystem zu einer höheren Benutzerzufriedenheit.
2. Die Adaption erhöht die Effizienz beim Auffinden gesuchter Informationen im Hilfesystem.
3. Eine benutzeradaptive Anpassung der Reihenfolge der Sprachoptionen fällt bei geeigneter Parameterwahl nicht negativ auf.
4. Die Erhöhung des Informationsgehalts der Sprachoptionen bei Experten wirkt sich positiv auf die Bedienung aus.
5. Die mehrstufig gegliederte Ausgabe von akustischer Information reduziert die mentale Belastung im Gegensatz zu einer ungegliederten Ausgabe.
6. Multimodale Interaktion führt zu effizienteren Dialogen hinsichtlich der Bediengeschwindigkeit sowie zu höherer Benutzerzufriedenheit.

7. Die zusätzliche visuelle Ausgabe der Hilfe forciert die Umstellung von Sprachbedienung auf haptische Bedienung.
8. Kontextunabhängige Hilfe wird in unbekannten Domänen gegenüber kontextabhängiger Hilfe bevorzugt.

7.2 Testdesign

Für die Durchführung der Tests stehen zwei Alternativen zur Wahl, das Between-Subject-Design (BSD) und das Within-Subject-Design (WSD) (Nielsen, 1997; Rubin, 1994). Beim BSD werden zwei verschiedene Produkte durch zwei unabhängige Testgruppen evaluiert, beim WSD werden zwei Produkte von einer Testgruppe evaluiert. Der Nachteil des WSD ist, dass die Versuchspersonen beim Test des ersten Produkts Wissen erwerben, das sich verfälschend auf die Testergebnisse für das zweite Produkt auswirken kann, da nicht mehr die gleichen Grundvoraussetzungen gegeben sind.

Im konkreten Fall bedeutet dies, dass Probanden bei der Verwendung von iDrive Wissen erwerben, welches das Verhalten im Umgang mit dem prototypischen Hilfesystem beeinflussen kann und umgekehrt. Insbesondere der Lerneffekt wirkt sich nachteilig auf die Durchführung der Tests aus, da zur Untersuchung der Hilfesysteme möglichst unerfahrene Benutzer benötigt werden. Aus diesem Grund wurde das BSD als Versuchsdesign gewählt.

7.3 Testplan

7.3.1 Versuchspersonen

Voraussetzung für die Teilnahme an den Tests war zum einen, dass die Probanden seit mindestens vier Jahren im Besitz eines Führerscheins sind oder alternativ eine Fahrleistung von 100.000 Kilometern aufweisen können, damit von einem stabilisierten Fahrverhalten ausgegangen werden kann. Zudem musste eine jährliche Fahrleistung von 10.000 Kilometern bestätigt werden, um die momentane Geübtheit zu gewährleisten (Salmen, 2003).

Des weiteren kamen nur interne BMW Mitarbeiter für die Teilnahme an den Versuchen in Frage, da der Prototyp auf Basis des iDrive Nachfolgers zum Zeitpunkt der Evaluierung der Geheimhaltung unterlag. Aus diesem Grund konnten besondere Kriterien wie etwa Alter oder Geschlecht bei der Auswahl der Probanden nicht berücksichtigt werden, da der Pool zur Verfügung stehender Versuchspersonen zu begrenzt war.

Insgesamt beteiligten sich 26 Personen (22 männlich, 5 weiblich) an den Versuchen, wobei 15 Personen auf den Test des Prototyps und 12 Versuchspersonen auf den iDrive Referenzversuch entfielen. Weitere Einzelheiten zum Profil der Probanden sind in Abbildung 7.1 ersichtlich.

VP-Nr.	Geschlecht	Alter	Schulabschluss	Beruf	Jahre mit Führerschein	km / Jahr	km / gesamt	Hand	BMW-Fahrer
VP 1-01	männlich	37	FH-Reihe	IT-Spezialist	19	10001-25000	>100000	rechts	ja
VP 1-02	männlich	31	Uni/FH	Diplom-Informatiker	13	10001-25000	>100000	rechts	ja
VP 1-04	männlich	43	Uni/FH	Dipl.-Ing. Elektrotechnik	25	10001-25000	>100000	rechts	ja
VP 1-06	männlich	29	Uni/FH	IT-Applikationsmanager	12	>25000	>100000	beide	ja
VP 1-07	männlich	31	Uni/FH	IT-Spezialist	13	10001-25000	>100000	rechts	ja
VP 1-08	männlich	35	Uni/FH	Entw. Ing.	19	>25000	>100000	links	ja
VP 1-09	männlich	29	Uni/FH	Informatiker	11	10001-25000	>100000	rechts	nein
VP 1-10	männlich	26	Abitur	PLV	8	10001-25000	>100000	rechts	nein
VP 1-11	männlich	30	Realschule	Prüfstandsfahrer	12	10001-25000	>100000	beide	nein
VP 1-13	männlich	20	Abitur	KFZ-Mechatroniker	2	10001-25000	0-100000	rechts	nein
VP 1-14	weiblich	25	Uni/FH	IT	7	0-10000	0-100000	rechts	nein
VP 1-15	männlich	35	Promotion	IT-Projektleiter	17	10001-25000	>100000	rechts	nein
VP 1-16	männlich	31	Uni/FH	Dipl. Ing / Doktorand	12	0-10000	>100000	rechts	nein
VP 1-17	weiblich	40	Promotion	Informatiker	22	>25000	>100000	rechts	ja
VP 1-18	weiblich	38	Uni/FH	Dipl.-Betriebswirt	20	10001-25000	>100000	rechts	ja
VP 2-01	männlich	24	Uni/FH	Student InfWiss	6	10001-25000	0-100000	rechts	nein
VP 2-02	männlich	26	Uni/FH	Student Winfo	8	>25000	>100000	rechts	nein
VP 2-03	weiblich	25	Uni/FH	Student InfWiss	7	0-10000	0-100000	rechts	nein
VP 2-04	männlich	27	Uni/FH	Student InfWiss	9	>25000	>100000	rechts	nein
VP 2-05	männlich	24	Uni/FH	Student InfWiss	6	0-10000	0-100000	rechts	nein
VP 2-06	männlich	36	Uni/FH	IT-Spezialist	18	0-10000	>100000	rechts	nein
VP 2-07	weiblich	37	Uni/FH	IT-Spezialist	19	10001-25000	>100000	rechts	ja
VP 2-08	männlich	23	Realschule	Fachinformatiker	5	10001-25000	>100000	rechts	ja
VP 2-09	männlich	28	Uni/FH	BWL	7	10001-25000	>100000	rechts	nein
VP 2-10	männlich	26	Uni/FH	IW	9	>25000	>100000	rechts	nein
VP 2-11	männlich	36	Uni/FH	Dipl. Ing.	18	>25000	>100000	rechts	ja
VP 2-12	männlich	36	Uni/FH	IT-Spezialist	18	>25000	>100000	rechts	ja

Abbildung 7.1: Übersicht über die Probanden

7.3.2 Versuchsaufbau

7.3.2.1 Versuchsaufbau im Fahrsimulator

Die Evaluierung des prototypischen Hilfesystems fand in einem Usabilitylabor bestehend aus einem Kontrollraum und einem statischen Fahrsimulator statt. Der Fahrsimulator besteht aus zwei Hauptkomponenten, einer Sitzkiste und einer Fahrsimulation. Die Sitzkiste entspricht hinsichtlich Fahrzeugführung, Armaturen, Bedienung, Instrumentierung und Anzeige-Bedienkonzept dem BMW 7er Nachfolger. Verschiedene Steuerungselemente vermitteln ein realistisches Fahrgefühl, z.B. Gegendruck am Lenkrad und Bremspedal. Beschleunigungs- oder Bremsverhalten können jedoch im statischen Simulator nicht nachempfunden werden.

Die Fahrsimulation wurde mit drei aufeinander abgestimmten Plasmabildschirmen dargestellt. Der Sichtbereich betrug ca. 130 Grad (vgl. Abbildung 7.2). Die

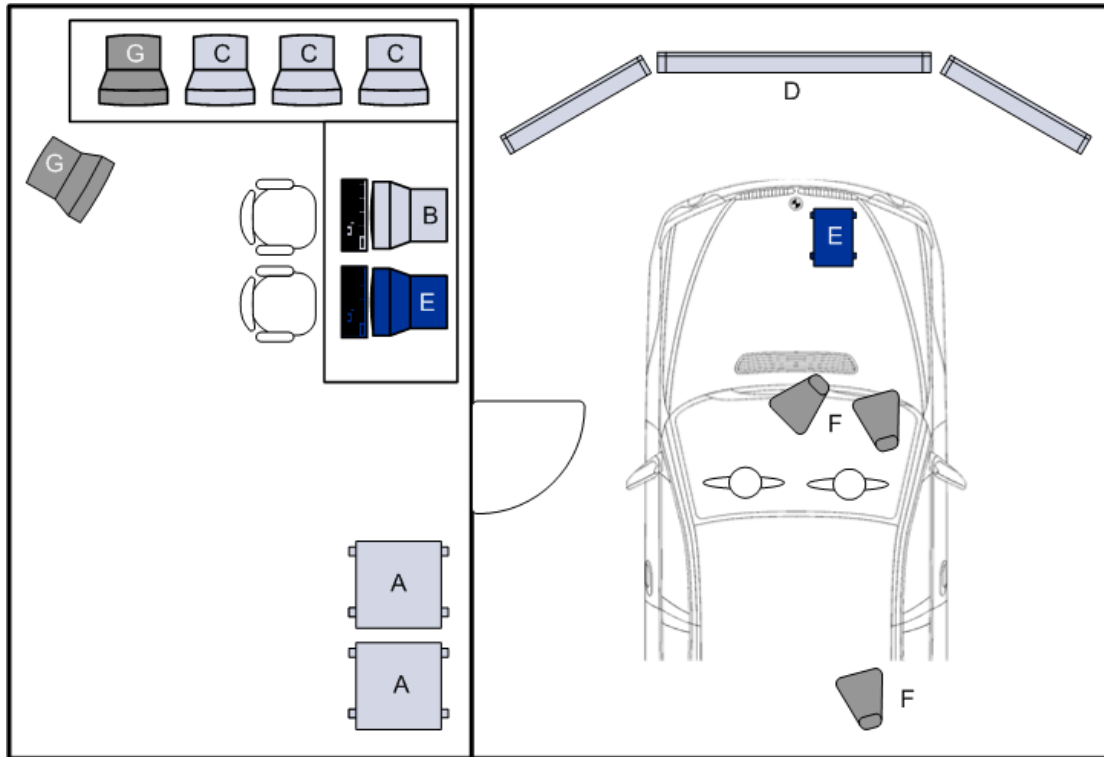


Abbildung 7.2: Versuchsaufbau Fahrsimulator: **A:** Server für die Fahrsimulation, **B:** Terminal zur Konfiguration der Fahrsimulation, **C:** Monitore zur Überwachung der Fahrsimulation (identisch zu Monitoren D), **D:** Plasmamonitore (linkes/mittleres/rechtes Sichtfeld), **E:** Rechner zur Steuerung der Dialogsimulation, **F:** Kameras, **G:** Videobearbeitung der Kamerabilder; Fahrer = Proband, Beifahrer = Versuchsleiter

Fahrsimulation stellte eine Autobahnfahrt dar, bei der die Aufgabe der Versuchsperson darin bestand, einem vorausfahrenden Fahrzeug in möglichst konstantem Abstand zu folgen. Die Geschwindigkeit des vorausfahrenden Fahrzeugs passte sich an auftretende Tempolimits an, ansonsten blieb die Geschwindigkeit konstant bei 120 km/h.

Interaktionen, Blickbewegungen und das Verkehrsgeschehen wurden mit drei Kameras aufgezeichnet. Jede Kamera lieferte ein Videosignal an einen Videobearbeitungsplatz im räumlich getrennten Kontrollbereich. Von hier aus erfolgte auch die Bedienung der Fahrsimulation und des Fahrerinformationssystems.

7.3.2.2 Versuchsaufbau im Fahrzeug

Die iDrive Referenzversuche fanden in einem Serien-BMW 3er statt. Dieser beinhaltete das aktuellste iDrive Bediensystem mit Professional Navigationssystem. Analog zum Aufbau im Fahrsimulator befinden sich drei Kameras im Innenraum, welche das Gesicht des Probanden, den Tacho und das Display (CID) aufzeichneten. Aus technischen und organisatorischen Gründen kamen im Fahrzeug jedoch lediglich Webcams zum Einsatz, weshalb aufgrund der wechselnden Lichtverhältnisse eine Aufzeichnung des Verkehrsgeschehens nicht möglich war.

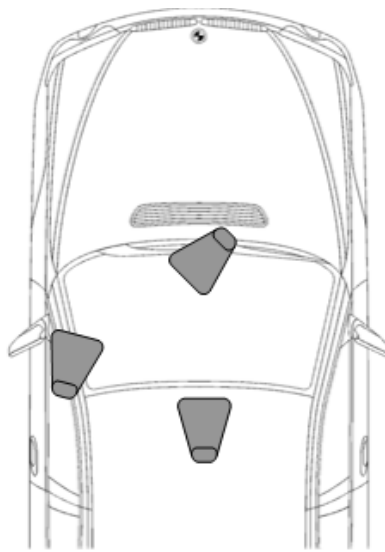


Abbildung 7.3: Versuchsaufbau im Fahrzeug

7.3.3 Versuchsablauf

Die Versuchsabläufe bei den Tests basierten auf Empfehlungen von Huber (2005) und Rubin (1994). Um einen stets gleichbleibenden Ablauf der Versuche zu gewährleisten, wurden die Tests mit einem Testdrehbuch durchgeführt. Dieses beschreibt den genauen Verlauf des Tests von der Begrüßung bis zur Verabschiedung des Probanden.

Die Versuche begannen mit einer Adaptionphase, in der die Probanden einen kurzen Überblick über den Versuchsaufbau erhielten. Darauf folgte die Instruktiionsphase, bei welcher der Testleiter eine formalisierte Einleitung zur Bedienung der Sitzkiste sowie zu Umfang und Zweck der Versuche gab (siehe Anhang C). Im Anschluß daran hatten die Probanden die Möglichkeit, ca. fünf bis sieben Minuten ohne weitere Instruktionen zu fahren. Dies ist sowohl im Fahrsimulator als auch

im Serienfahrzeug wichtig, um negative Einflüsse durch ungewohnte Situationen zu minimieren. Im weiteren Verlauf wurden die einzelnen Aufgaben bearbeitet (ca. 60 Minuten incl. 5 Minuten Pause).

Nach Beendigung der letzten Aufgabe wurde der Fahrsimulationsraum verlassen bzw. im iDrive Referenztest der nächste Parkplatz angefahren. Die Nachbereitung der Versuchsfahrten fand nach einer kurzen Pause statt. Die Probanden erhielten zu diesem Zweck einen Fragebogen, den sie ausfüllen mussten (ca. 20 Minuten) (siehe Anhang D).

7.3.4 Aufgabenstellung

Die Aufgaben bestanden aus einem Querschnitt von Funktionen aus verschiedenen Systembereichen. Bei der Aufgabenstellung wurde darauf geachtet, dass möglichst keine Schlüsselwörter oder Shortcuts in der Formulierung auftauchten. Folgende Aufgaben waren von den Probanden zu bearbeiten:

1. Teilnehmer anrufen: Sie möchten Markus Abel aus Ihrem Telefonbuch anrufen. Im Telefonbuch angekommen, können Sie durch das Kommando Sprachoptionen herausfinden, welche Möglichkeiten Sie haben, diesen Anruf zu erledigen.
2. Wahlwiederholung: [*Eingabe einer beliebigen Telefonnummer durch den Testleiter*] Sie möchten die Nummer während der Fahrt nicht nochmals eintippen. Finden Sie deshalb eine andere Möglichkeit, die eben gewählte Nummer nochmals anzurufen.
3. Nachricht löschen: Im Nachrichtenspeicher befindet sich eine SMS mit einer unbekannten Nummer; also kein Name aus dem Adressbuch sondern nur eine Ziffernfolge. Sie möchten nun diese SMS löschen.
4. Kartenansicht wechseln: Sie möchten die verschiedenen Möglichkeiten der Kartenansicht kennen lernen. Dazu müssen Sie zunächst herausfinden, wo Sie die Darstellung der Karte ändern können.
Wenn Sie das Menü gefunden haben, möchten Sie die Karte in Fahrtrichtung ausgerichtet sehen.
5. Kartenmaßstab wählen: Auch der Maßstab der Karte ist einstellbar. Bitte finden Sie heraus, wie das funktioniert.
6. Routenkriterien verändern: Sie wollen die Streckenführung beeinflussen, also beispielsweise ein Kriterium wie die Benutzung von Autobahnen. Finden Sie das dazu nötige Menü.
7. Kontextunabhängige Hilfe: Sie möchten den Sender Bayern 3 hören. Bitte finden Sie mit einer Frage den Weg zum Ziel.
8. Verkehrsfunk aktivieren: Sie möchten den Verkehrsfunk anstellen.

9. Verkehrsfunk deaktivieren: Sie möchten den Verkehrsfunk wieder deaktivieren.
10. CD und CD-Titel auswählen: Sie möchten eine CD hören. Wählen Sie Track 5 von CD 3 im CD-Wechsler.
11. CD-Titel auswählen: Nun möchten Sie zu Eintrag 8 auf derselben CD springen.
12. Vorheriger/Nächster Titel: Sie möchten die Titel der Reihe nach hören. Spielen Sie einen Titel nach dem anderen ab.
13. Klangeinstellungen: Sie möchten nun wiederum durch eine Frage erfahren, wo Sie die Bässe verstellen können. Bitte fragen Sie nach der Möglichkeit der Bässeverstellung.

7.4 Ergebnisse

Die Auswertung der Versuche erfolgt getrennt nach den Versuchsgruppen im Fahrsimulator (V1) und im Fahrzeug (V2). Des weiteren wird die Versuchsgruppe V1 in zwei Untergruppen aufgeteilt. Dies scheint sinnvoll, da im Rahmen der Versuche mit dem Prototyp wie in Kapitel 6.1 beschrieben ein neuer Spracherkenner zum Einsatz kam. Die Erkennerleistung des Spracherkenners war im Rahmen von Vortests und zu Beginn der Versuchreihe mäßig. Erst im weiteren Verlauf der Untersuchung konnte die Erkennerleistung optimiert werden. Da jedoch die Versuchspersonen das Testsystem als Ganzes wahrnehmen und sich die schlechte Erkennerleistung auf die Gesamtbewertung des Prototyps auswirkte, wurden die Versuchsgruppe je nach Erkennerleistung in entsprechende Untergruppen aufgeteilt (siehe Abbildung 7.4).

System	Prototyp		Referenzsystem
Testort	Fahrsimulator		Fahrzeug
Probandenbezeichnung	VP1_xx		VP2_xx
Gruppe	V1a	V1b	V2
Erkennungserleistung	mäßig	gut	gut
Probandenzahl	13	5	12

Abbildung 7.4: Gruppierung der Probanden

7.4.1 Bewertung allgemeiner Systeminteraktionen

Da im Rahmen der Untersuchung im Fahrsimulator ein neuer Spracherkenner mit einer anfänglich schlechteren Erkennungserleistung eingesetzt wurde, weisen einige Be-

obachtungen, die in engem Zusammenhang mit der Erkennerleistung stehen, tendenziell schlechtere Ergebnisse auf. Die Analyse der Erkennerfehler zeigt, dass V1a z.T. erhebliche Einbußen in der Erkennrate verzeichnete (siehe Abbildung 7.5). Vor allem bei den ersten Aufgaben in der einzelnen Versuchen war die Zahl der Erkennerfehler sehr hoch, da die Probanden noch nicht an das SDS gewöhnt waren. Sprachlautstärke, Sprachgeschwindigkeit und Prosodie wurden erst im weiteren Versuchverlauf optimal an das SDS angepasst. Zudem produzierten Nebengeräusche (Räuspern, Murmeln, etc.) ebenfalls unerwünschte Spracheingaben. Die Erkennrate bei V1b dagegen bewegt sich in etwa auf dem Niveau von V2.

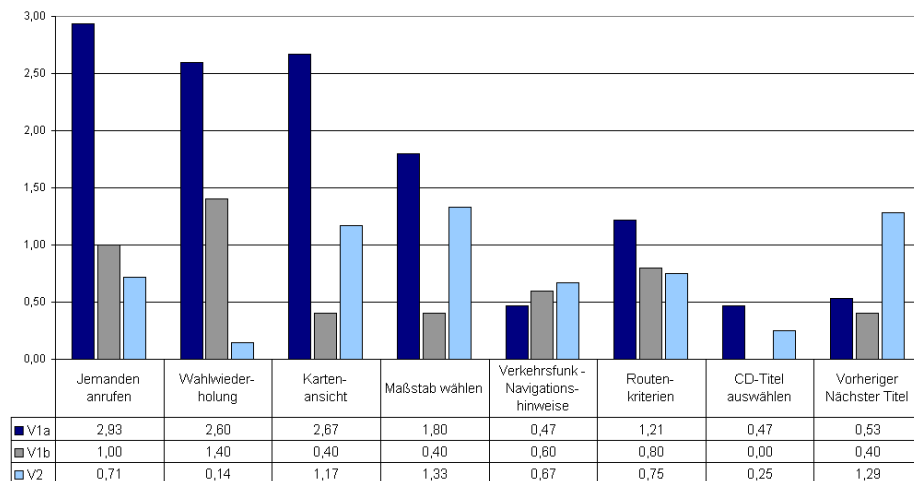


Abbildung 7.5: Erkennerfehler (Mittelwert)

Als Maß zur Bestimmung der Effizienz wird die Aufgabendauer herangezogen. Abbildung 7.6 zeigt die Zeit, die die Probanden benötigten, um eine Aufgabe zu bearbeiten. Die Analyse der Ergebnisse zeigt, dass V1b stets die kürzesten Bearbeitungszeiten aufweist. V1a ist zumeist mit längeren Bearbeitungszeiten verbunden, was jedoch auf die schlechtere Erkennerleistung zurückzuführen ist. Einzelne Kommandos mussten oft mehrmals wiederholt werden, um das gewünschte Dialogziel zu erreichen. V2 geht trotz guter Erkennerleistung stets mit längeren Bearbeitungszeiten als V1b einher, in sechs Fällen benötigten die Probanden sogar länger als in V1a. Die Bearbeitung der Aufgaben verläuft somit beim Prototyp deutlich effizienter.

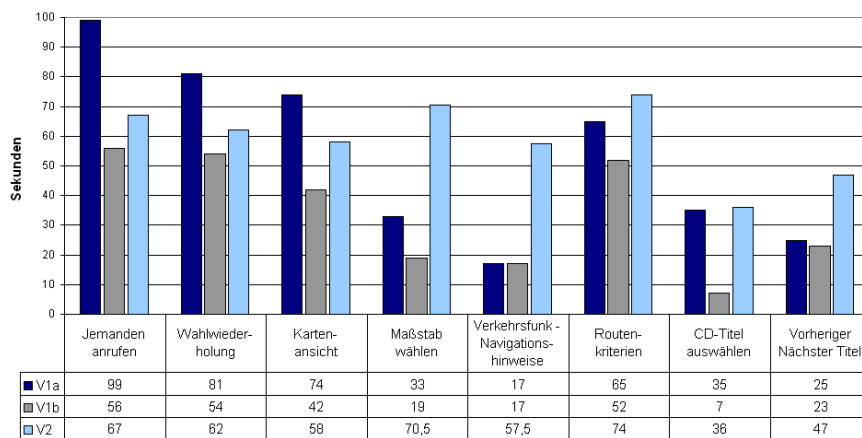


Abbildung 7.6: Aufgabendauer (Median)

Neben der Aufgabendauer gibt auch die Anzahl der Sprecherbeiträge Aufschluß über die Effizienz des Systems (siehe Abbildung 7.7). Auch hier weist V1b sehr gute Resultate auf, in fünf Aufgaben sogar die besten. Die höheren Beiträge bei V1a sind wiederum auf die schlechte Erkennerleistung zurückzuführen. Das iDrive Referenzsystem (V2) erfordert trotz sehr guter Erkennerleistung meist mehr Sprecherbeiträge als der Prototyp (V1).

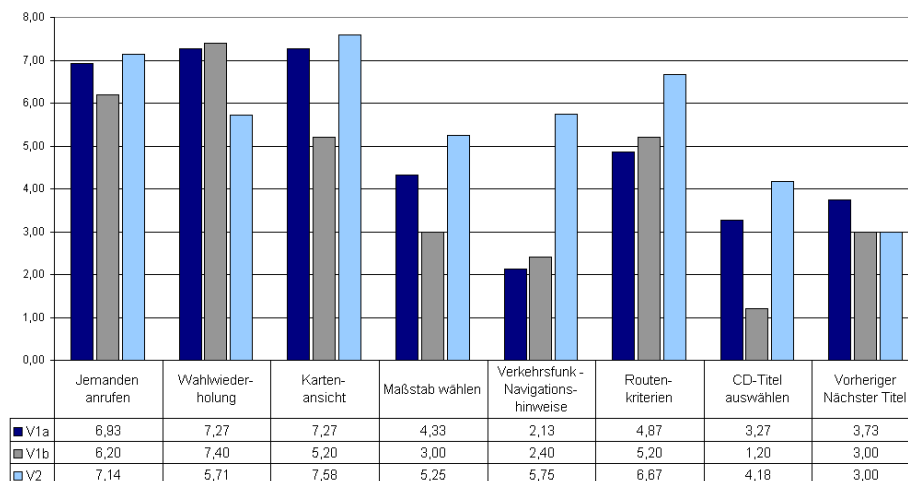


Abbildung 7.7: Sprecherbeiträge (Mittelwert)

Als möglicher Grund kann die Aufteilung der Optionen in verschiedene Ebenen angeführt werden. Falls im iDrive SDS die Optionen lang ausfallen oder ein ge-

suchtes Kommando in der Mitte einer längeren Sprachausgabe liegt, kann dies zu einer schlechten Memorierung und zu wiederholtem Aufruf der Hilfe führen. Diese These kann durch die Betrachtung der Optionsaufrufe gestützt werden (siehe Abbildung 7.8). Mit Ausnahme der letzten Aufgabe wurden die Optionen bei V2 häufiger aufgerufen als bei V1a oder V1b. Im internen Vergleich für V1 liefert V1b erneut die tendenziell besseren Resultate.

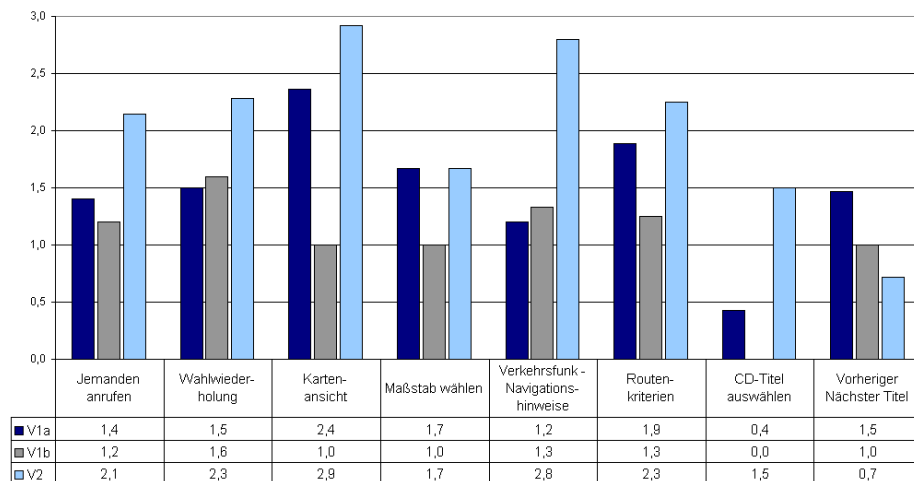


Abbildung 7.8: Optionsaufrufe (Mittelwert)

7.4.2 Analyse des Blickverhaltens

Das Blickverhalten der Versuchspersonen wurde anhand der Videoaufzeichnungen analysiert. Diese beinhalteten im Fahrsimulator Informationen zu Blicken auf das CID, den Tachobereich und den Verkehr. Die restlichen Blicke wurden der Menge „Sonstige Blicke“ zugeordnet. Eine Auswertung der Spiegelblicke war nicht möglich, da im Fahrsimulator keine Innen- oder Außenspiegel vorhanden waren. Die Videos wurden mit Hilfe der Software Interact framegenau ausgewertet.

Von Interesse sind bei der Auswertung Fixationen auf Bereiche im Fahrzeuginnenen, die eine Verweildauer von mehr als zwei Sekunden besitzen. Nach Schweigert (2002) sollte der Anteil an Blicken, die nicht auf das Verkehrsgeschehen gerichtet sind und eine Fixationsdauer von mehr als zwei Sekunden besitzen, 5% der gesamten Blicke nicht übersteigen. Sowohl bei V1 als auch bei V2 wird dieser Wert nicht überschritten (siehe Abbildung 7.9, Zeile 19).

#	Parameter	V1a	V1b	V2
1	Blicke pro Minute (Blickwechsel)	51,92	46,67	28,46
2	Blicke Display pro Minute	17,14	14,74	7,75
3	Blicke Verkehr pro Minute	23,60	22,21	13,78
4	Blicke Tacho pro Minute	10,48	9,09	3,78
5	Blicke VL pro Minute	0,19	0,09	0,43
6	Blicke Sonstiges pro Minute	0,50	0,55	2,72
7	Mittel Blickdauer Gesamt [in Sekunden]	1,20	1,32	2,22
8	Mittel Blickdauer Display [in Sekunden]	0,96	0,96	1,03
9	Mittel Blickdauer Verkehr [in Sekunden]	1,64	1,86	3,63
10	Mittel Blickdauer Tacho [in Sekunden]	0,61	0,59	0,57
11	Mittel Blickdauer VL [in Sekunden]	0,67	0,81	0,93
12	Mittel Blickdauer Sonstiges [in Sekunden]	0,53	0,57	0,89
13	Verhältnis Blicke Gesamt über 2s / Blicke Gesamt	0,12	0,14	0,25
14	Verhältnis Blicke Display über 2s / Blicke Display	0,05	0,04	0,07
15	Verhältnis Blicke Verkehr über 2s / Blicke Verkehr	0,22	0,26	0,46
16	Verhältnis Blicke Tacho über 2s / Blicke Tacho	0,01	0,01	0,01
17	Verhältnis Blicke VL über 2s / Blicke VL	0,02	0,00	0,06
18	Verhältnis Blicke Sonstiges über 2s / Blicke Sonstiges	0,01	0,03	0,07
19	Verhältnis Blicke Display+Tacho über 2s / Blicke Gesamt	0,02	0,01	0,02
20	Verhältnis Blicke Display+Tacho über 2s / Blicke Display+Tacho	0,03	0,03	0,05
21	Anteil Blicke Display an Blicke Gesamt	0,33	0,31	0,27
22	Anteil Blicke Verkehr an Blicke Gesamt	0,46	0,48	0,48
23	Anteil Blicke Tacho an Blicke Gesamt	0,20	0,20	0,14
24	Anteil Blicke VL an Blicke Gesamt	0,00	0,00	0,01
25	Anteil Blicke Sonstiges an Blicke Gesamt	0,01	0,01	0,10

Abbildung 7.9: Analyse des Blickverhaltens

In den Versuchen V1a und V1b fanden häufigere Blickwechsel statt als bei V2, d.h. die Probanden wechselten häufiger zwischen Verkehr und Display, Tacho und Sonstigem (siehe Abbildung 7.9, Zeile 1). Die verstärkten Blicke auf den Tacho lassen sich im Fahrsimulator durch ein fehlendes physisches Geschwindigkeitsempfinden erklären (siehe Abbildung 7.9, Zeile 4). Die verstärkten Blicke auf das CID bei V1a und V1b können durch die visuelle Komponente des Hilfesystems erklärt werden, die bei V2 nicht vorhanden war (siehe Abbildung 7.9, Zeile 2). Für den Anteil an Displayblicken zu den Gesamtblicken liefern V1a, V1b und V2 keine signifikanten Unterschiede, alle drei Werte bewegen sich ca. 30%, tendenziell liefert jedoch V2 erwartungsgemäß mit 27% den besten Wert.

7.4.3 Subjektive Bewertung der Hilfesysteme

Neben objektiven Messungen liefern subjektive Eindrücke der Probanden ein zusätzliches Bild über das prototypische Hilfesystem. Nachfolgend werden die zen-

tralen Ergebnisse der Fragebogenauswertung erläutert (siehe Anhang D). Eine umfassende Auswertung des Fragebogens ist in Sagstetter (2007) zu finden.

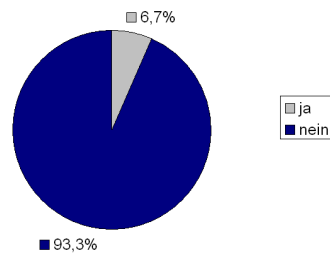


Abbildung 7.10: Ist Ihnen im Laufe des Tests eine Veränderung hinsichtlich der Länge oder der Inhalte der Hilfe aufgefallen?

80% der Probanden aus Versuchreihe V1 bewerteten die individuelle Anpassung der Hilfeinhalte an Wissen und Präferenzen eines Benutzers überwiegend positiv. Neben der Bewertung auf einer Skala hatten die Versuchspersonen auch die Möglichkeit, eine Begründung für ihre Einschätzung zu geben. Zusammenfassend lässt sich festhalten, dass sich die Probanden durch die Adaption einfachere und kürzere Dialoge sowie eine geringere mentale Belastung erwarten. Die mit einem adaptiven Konzept einhergehende Inkonsistenz der Dialoge wurde von den Probanden nicht bemerkt bzw. nicht als negativ empfunden (siehe Abbildung 7.10).

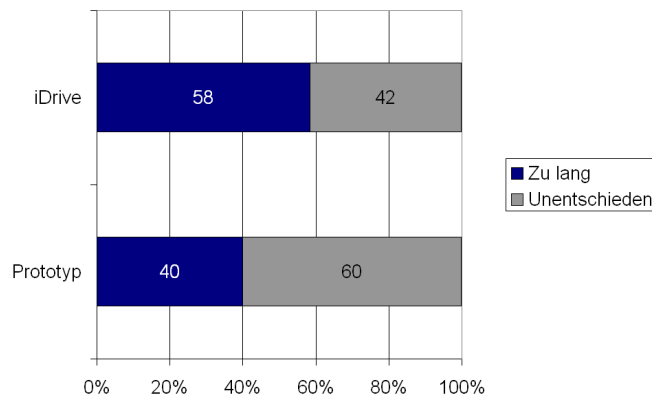


Abbildung 7.11: Wie beurteilen Sie die Länge der Hilfeausgaben?

Die Länge der Hilfeausgaben wurde bei V2 durch 58% der Probanden als zu lang eingestuft, bei V1 geschah dies nur in 40% der Fälle (siehe Abbildung 7.11). Neben der Länge der Ausgaben wurde auch der Informationsgehalt bewertet. Dazu wurden die Probanden befragt, ob nach Verwendung der Hilfe noch Unklarheiten darüber bestanden, welche Funktion zur Lösung einer Aufgabe zur Verfügung steht bzw. wie diese Funktion angewendet wird.

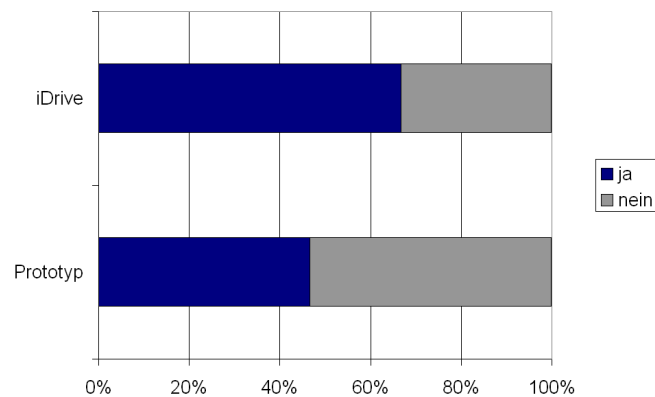


Abbildung 7.12: Das Kommando „Sprachoptionen“ wird immer dann benutzt, wenn Befehle in einem Kontext nicht bekannt sind. Traten bei Ihnen nach der Benutzung des Kommandos „Sprachoptionen“ noch Unklarheiten auf?

67% der Probanden bei V2 antworteten, dass nach Verwendung der Hilfe z.T. noch Unklarheiten auftraten. Im Gegensatz dazu traten bei V1 nur bei 47% der Probanden Unklarheiten auf (siehe Abbildung 7.12). Die visuelle Unterstützung während der akustischen Hilfeausgaben wurde von 73% der Probanden als nützlich empfunden (siehe Abbildung 7.13). Aus den Kommentaren der Versuchspersonen geht hervor, dass die zeitlich synchrone Hervorhebung der Sprachkommandos auf dem CID die Orientierung im Dialog erleichterte.

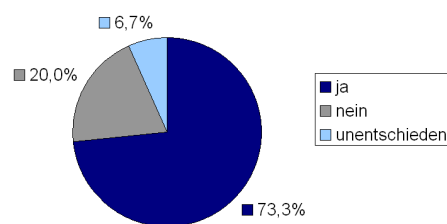


Abbildung 7.13: Empfanden Sie die zusätzliche visuelle Anzeige der Hilfeausgaben als nützlich?

Ogleich nur bei 2% aller Interaktionen der Controller zur Auswahl eines Sprachkommandos verwendet wurde, schätzten 60% der Probanden die Möglichkeit multimodaler Interaktionen als sehr nützlich ein (siehe Abbildung 7.14). Dennoch übertraf die Barge-In Funktion mit 9% aller Interaktionen die Push-In Funktion deutlich.

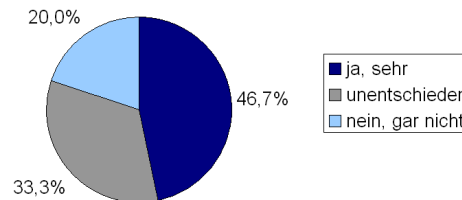


Abbildung 7.14: Die Auswahl der Sprachoptionen kann sowohl durch Sprache als auch durch den Controller erfolgen. War diese Möglichkeit für Sie persönlich von Nutzen?

Die kontextunabhängige Hilfe wurde insgesamt ebenfalls positiv bewertet, 67% der Probanden aus V1 bewerteten diese Form der Unterstützung als hilfreich. Abbildung 7.15 zeigt zudem, dass 73% der Versuchspersonen die kontextunabhängige Hilfe mit der Note 1 bewerteten (1=sehr nützlich, 6=gar nicht nützlich).

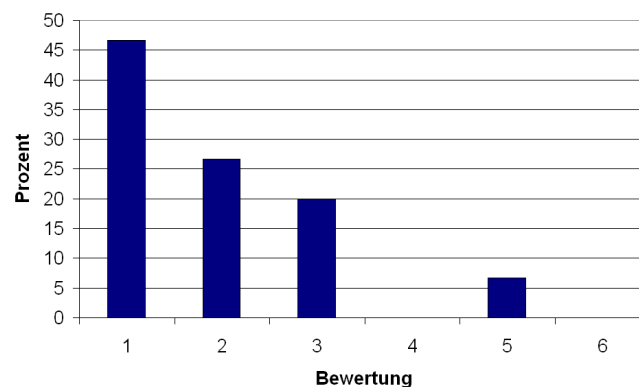


Abbildung 7.15: Wie nützlich ist für Sie diese Art der [kontextunabhängigen] Hilfe?

7.4.4 Ermittlung der Usability anhand des Evaluierungsframeworks PARADISE

7.4.4.1 Beschreibung von PARADISE

Das von Walker et. al. (1997) entwickelte PARADISE (Paradigm for Dialogue System Evaluation) Evaluierungsframework bündelt verschiedene Ansätze zur Be-

wertung von SDS. Walker et. al. geht davon aus, dass ein Zusammenhang zwischen der objektiv messbaren Leistung und der subjektiven Erfahrung eines Benutzers im Umgang mit einem SDS besteht. PARADISE kombiniert verschiedene Performanzmaße und ermöglicht so eine umfassende Bewertung des gesamten SDS.

In Anlehnung an Walker et. al. (1997) ist das oberste Ziel eines SDS, die Nutzerzufriedenheit zu maximieren. Dazu muss zum einen der Aufgabenerfolg maximiert werden und zum anderen müssen die Dialogkosten bestehend aus Effizienzmaßen und qualitativen Maßen minimiert werden (vgl. Abbildung 7.16). Die Vorhersage

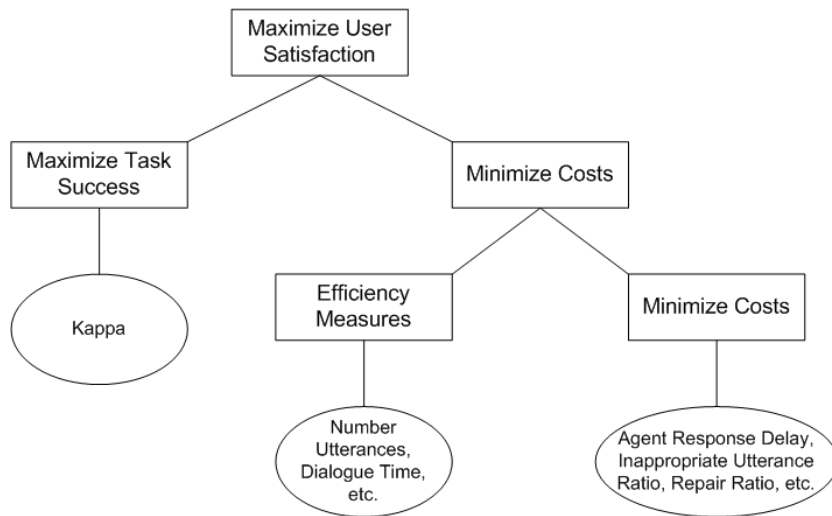


Abbildung 7.16: Grundlagen des PARADISE Frameworks. Quelle: Walker et. al. (1997)

der Performanz eines SDS erfolgt mittels einer multivariaten linearen Regressionsanalyse. Diese basiert auf der Summe des Maßes für den Aufgabenerfolg (κ) und den Dialogkosten (c_i). Die Parameter α und w_i dienen zur Gewichtung der jeweiligen Performanzkomponenten.

$$Performance = (\alpha \cdot N(\kappa)) - \sum_{i=1}^n w_i \cdot N(c_i) \quad (7.1)$$

Die Berechnung von κ setzt eine Notation der Interaktionen eines Benutzers mit dem SDS in Form einer Attribut-Wert-Matrix (AVM: Attribute-Value-Matrix) voraus. Eine AVM in einem Zugauskunftssystem kann z.B. der Abfahrtsbahnhof ein Attribut, das verschiedene Städte als Werte annehmen kann. Die Beschreibung einer Aufgabe in einem SDS erfolgt durch Instanziierung der AVM in einer Konfusionsmatrix. In die erste Zeile werden die Attributwerte (Schlüssel) eingetragen, die zwischen SDS und Benutzer ausgetauscht werden müssen, um eine Aufgabe vollständig zu erledigen. In die erste Spalte der Konfusionsmatrix werden die

Attribute	Possible values	Information flow
depart-city (DC)	Milano, Roma, Torino, Trento	to agent
arrival-city (AC)	Milano, Roma, Torino, Trento	to agent
depart-range (DR)	morning, evening	to agent
depart-time (DT)	6am, 8am, 6pm, 8pm	to user

Tabelle 7.1: Attribut-Wert-Matrix. Quelle: Walker et. al. (1997)

Attributwerte eingetragen, die tatsächlich ausgetauscht wurden (siehe Abbildung 7.17). Wenn eine Informationseinheit zu einem Schlüssel passt, wird der Wert in

	KEY													
	DEPART-CITY				ARRIVAL-CITY				DEPART-RANGE		DEPART-TIME			
DATA	v1	v2	v3	v4	v5	v6	v7	v8	v9	v10	v11	v12	v13	v14
v1	22		1		3									
v2		29												
v3	4		16	4			1							
v4	1	1	5	11			1							
v5	3				20									
v6						22								
v7			2		1	1	20	5						
v8			1		1	2	8	15						
v9									45	10				
v10									5	40				
v11											20		2	
v12											1	19	2	4
v13											2		18	
v14											2	6	3	21
sum	30	30	25	15	25	25	30	20	50	50	25	25	25	25

Abbildung 7.17: Konfusionsmatrix: In der ersten Zeile werden die notwendigen Attributwerte eingetragen, in der ersten Spalte die tatsächlich ausgetauschten. $v1$ bis $v4$ repräsentieren hier die vier möglichen Abfahrtsbahnhöfe, $v5$ bis $v8$ die möglichen Zielbahnhöfe etc. Quelle: Walker et. al. (1997)

der zugehörigen Diagonalzelle um den Wert 1 erhöht. Zellen außerhalb der Diagonale repräsentieren Missverständnisse, die während eines Dialogs nicht korrigiert werden.

Der Aufgabenerfolg κ wird mit folgender Formel berechnet:

$$\kappa = \frac{P(A) - P(E)}{1 - P(E)} \quad (7.2)$$

$P(A)$ beschreibt hierbei die Übereinstimmungshäufigkeit zwischen den tatsächlichen Dialogwerten und den Schlüsseln:

$$P(A) = \frac{\sum_{i=1}^n M(i, i)}{T} \quad (7.3)$$

$M(i, i)$ beschreibt die Werte der Diagonalen in der Konfusionsmatrix M , T repräsentiert die Summe der auftretenden Schlüssel ($M_{t_1} + \dots + M_{t_n}$; bezogen auf das Beispiel aus Abbildung 7.17 beträgt $T = 400$). $P(E)$ beschreibt die Häufigkeit, mit der erwartet wird, dass die Werte aus der AVM und die Schlüssel zufällig übereinstimmen:

$$P(E) = \sum_{i=1}^n \left(\frac{t_i}{T}\right)^2 \quad (7.4)$$

t_i repräsentiert die Summe der Häufigkeiten in Spalte i in M .

Die Berechnung der Dialogkosten $\sum_{i=1}^n w_i \cdot N(c_i)$ der Performanzfunktion (Gleichung 7.1) ist aufgeteilt in die Bereiche Dialogeffizienz und Dialogqualität. Maße zur Bestimmung der Effizienz sind Aufgabendauer, Systembeiträge und Nutzerbeiträge innerhalb eines Dialogs. Als Qualitätsmaße können die Erkennungsqualität, die Anzahl an Systemausgaben, Timeouts, Erkennenfehler, Hilfeanfragen und Barge-Ins herangezogen werden. Jeder dieser Faktoren wird als eine Funktion c_i dargestellt und entsprechend seiner Bedeutung mit w_i gewichtet.

7.4.4.2 Kritik an PARADISE

Der Einsatz von PARADISE zur Bewertung von Sprachdialogen ist nicht unumstritten. Ein Kritikpunkt betrifft die zur Berechnung von κ notwendige Notation der Dialoge in einer AVM. Die Notation der Dialoge setzt voraus, dass die Menge an auszutauschenden Informationen stets gleich bleibt. Im Falle der oben dargestellten Zugauskunft trifft dies zu, wenn jedoch die ausgetauschten Informationen bei verschiedenen Benutzern variieren (vgl. Hjalmarsson (2002)), liefert PARADISE verfälschte Informationen, da verschieden komplexe Dialoge auch unterschiedliche Dialogkosten verursachen.

Ein weiterer Kritikpunkt betrifft die Annahme, dass Systemperformanz und Benutzerzufriedenheit linear miteinander korrelieren. Nach Larsen (2003) existieren keine Beweise, die diese Annahme stützen würden. Es gebe zwar einen Zusammenhang zwischen Spracherkennungsqualität und Benutzerzufriedenheit, dieser sei jedoch selten linear.

7.4.4.3 Anpassung der Methodik nach Hassel

In Hassel (2006) wird PARADISE an die Rahmenbedingungen bei der Untersuchung von Fahrerinformationssystemen angepasst. Nachfolgend werden die zentralen Aspekte der veränderten Methodik erläutert, für eine ausführliche Darstellung der Thematik sein an Hassel (2006) und Sagstetter (2007) verwiesen.

Die Bestimmung der Nutzerzufriedenheit erfolgt mit Hilfe eines Fragebogens, der jeweils am Ende einer Versuchreihe ausgefüllt wird. Daraus lassen sich drei Maße zur Beschreibung der Nutzerzufriedenheit ableiten (siehe Tabelle 7.2). Die Bewer-

Wert	Fragebogen
US	Das System war sehr benutzerfreundlich.
US1	Ich konnte alle Aufgaben ohne Probleme bewältigen. Das Sprachdialogsystem war sehr benutzerfreundlich. Ich habe mich während der Bedienung sehr geärgert.
US2	Die Sprachsteuerung im Fahrzeug ist sehr nützlich. Ich würde die Sprachsteuerung weiterempfehlen. Ich möchte gerne eine Sprachsteuerung im Fahrzeug haben.

Tabelle 7.2: Beschreibung der Nutzerzufriedenheit nach Hassel (2006)

tung der einzelnen Aussagen geschieht auf einer Skala von 1 (schlechtestes Ergebnis) und 4 (bestes Ergebnis). Die Zufriedenheitsmaße werden durch die Bildung des Mittelwerts der einzelnen Bewertungen bestimmt.

Der Aufgabenerfolg wird mit einem veränderten Maß κ^* dargestellt.

κ^* spiegelt die zufallsbereinigte Übereinstimmung zwischen System und Nutzer wider. Je höher die κ^* -Werte, desto besser [ist] die Übereinstimmung zwischen dem, was der Benutzer sagt, und dem, was das System versteht. (Hassel, 2006)

Die Berechnung von κ^* basiert zudem nicht mehr auf einer AVM, sondern auf einem gerichteten, verbundenen Graphen. Dieser zeigt die möglichen Lösungswege auf, die ein Benutzer zur Erfüllung einer Aufgabe verwenden kann. Die resultierende Konfusionsmatrix ist in Abbildung 7.18 dargestellt. Die erste Spalte

	Hauptmenu	Kommunikation	Telefon	Telefonbuch	AZ Liste	Name wählen	Markus Abel	Ja	(Sprach-)Optionen	Zurück	Abbruch	FALSCH	SUMMEN
Hauptmenu													0
Kommunikation													0
Telefon			2									1	2
Telefonbuch													0
AZ Liste													0
Name wählen													0
Markus Abel							1					3	1
JA								1					1
(Sprach-)Optionen									1				1
Zurück										1			0
Abbruch											1		0
FALSCH												1	0
SUMMEN	0	0	2	0	0	0	1	1	1	0	0	4	

Abbildung 7.18: Konfusionsmatrix nach Hassel (2006)

der Konfusionsmatrix enthält die Nutzeräußerungen, die erste Zeile beinhaltet die vom System verstandenen Kommandos. Zusätzlich zu den zur Aufgabenlösung notwendigen Sprachbefehlen wurden die Kommandos „ja“, „Optionen“, „zurück“

und „Abbruch“ aufgenommen. Des weiteren stellt die Konfusionsmatrix Mechanismen bereit, um fehlerhafte Eingaben abbilden zu können. Im Gegensatz zum ursprünglichen Ansatz bei PARADISE werden hier alle Beiträge eines Benutzers zur Bewertung des Aufgabenerfolgs κ^* herangezogen. Die Spalte „FALSCH“ repräsentiert alle Eingaben, bei denen ein Benutzer ein Kommando richtig verwendete, das SDS jedoch das Falsche verstand. Die Zeile „FALSCH“ umfasst ungültige Benutzereingaben, die vom System richtig bzw. im Schnittpunkt FALSCH/FALSCH korrekt falsch interpretiert wurden.

Hassel führt weiterhin das Erfolgsmaß E ein. Dieses Maß spiegelt wider, wieviel ein Benutzer bzw. das SDS zum Erreichen eines Dialogziels beitragen. E kann die Werte -1 (alle Beiträge führten den Dialog weg vom ursprünglichen Ziel), 0 (kein Einfluss) oder 1 (alle Beiträge führten zum Ziel hin) annehmen (siehe Tabelle 7.3). Die Bewertung von E erfolgt getrennt für das SDS und den Benutzer.

7.4.4.4 Auswertung der Versuchreihen

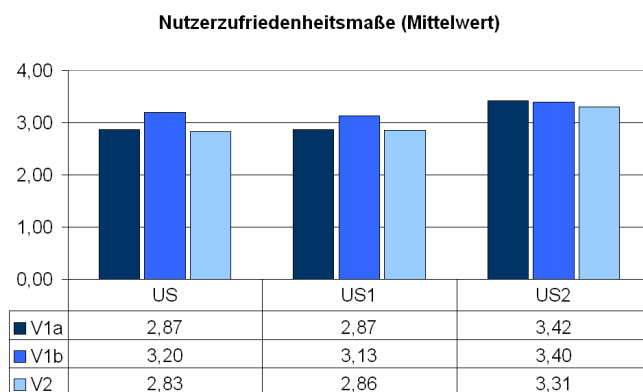


Abbildung 7.19: Nutzerzufriedenheitsmaße

Die Nutzerzufriedenheit US (siehe Abbildung 7.19) unterscheidet sich bei den Versuchen V1a und V2 nahezu gar nicht. Der Prototyp mit guter Spracherkennung (V1b) hingegen liefert einen Durchschnittswert von 3,20 und liegt damit deutlich im Bereich „sehr zufrieden“. Ein ähnliches Resultat liefert auch die Auswertung von US1. Während V1a (2,87) und V2 (2,86) nahezu identische Werte besitzen, spiegelt V1b mit einer

Bewertung von 3,13 eine deutlichere Zufriedenheit der Probanden mit dem SDS wider. Bei den Bewertungen für US2 sind kaum Unterschiede zu verzeichnen, der Prototyp erhält mit V1a und V1b jedoch eine geringfügig bessere Bewertung als das iDrive System in V2. Tendenziell wird der Prototyp stets besser bewertet als das Referenzsystem. Zieht man lediglich V1b und V2 zum Vergleich heran, schneidet der Prototyp bei gleicher Erkennerleistung bei den bedeutenderen Dimensionen US und US1 wesentlich besser ab. Selbst bei schlechter Erkennerleistung schneidet V1b noch geringfügig besser als ab V2.

Die Bewertung hinsichtlich des Aufgabenerfolgs κ^* offenbart die negativen Einflüsse der schlechten Spracherkennungsleistung bei V1a. Insofern liefert V1a mit Ausnahme von zwei Aufgaben tendenziell die schlechtesten Werte (siehe Abbil-

Nutzerbeiträge		
Parameter	Wert	Erläuterung
OK	1	Richtiger Nutzerbeitrag
Autokorrektur	1	Nutzer bemerkt einen Fehler und korrigiert ihn sofort
Wordspotting	1	Richtiger Beitrag eingebettet in weiteren Phrasen
Modell	-1	Der Beitrag geht auf ein falsches Modell über das System zurück
OOT	-1	Der Beitrag wird vom Erkennen verstanden, aber er bezieht sich auf eine andere Aufgabe
OOV	-1	nutzer verwendet in Erkennerswortschatz nicht vorhandene Synonyme von gültigen Befehlen
Noise	-1	Benutzer spricht nur unverständliches
Abbruch	-1	Benutzer bricht den Dialog bewusst ab
Barge-In	0	Barge-In
Optionen	0	Aufruf der Hilfe
Systembeiträge		
Parameter	Wert	Erläuterung
OK	1	Richtiger Nutzerbeitrag
Richtig negativ	1	Unverständliche Eingabe als solche erkannt
Folgerichtig	1	Nutzerbeitrag führt in einen positiven, nicht beabsichtigten Zustand
Disambi	1	Unverständliche Eingabe erkannt + Reparaturdialog
Nicht folgerichtig	-1	Richtiger Nutzerbeitrag führt in einen negativen Zustand
OOT	-1	OOT-Nutzerbeitrag wird richtig verstanden
Attribut	-1	Befehl wird nur zum Teil richtig verstanden
Falsch (negativ)	-1	Richtiger Beitrag wird nicht erkannt

Tabelle 7.3: Kriterien zur Berechnung des Erfolgsmaßes E . Quelle: Hassel (2006) und Sagstetter (2007)

dung 7.20). V1b und V2 erreichen jeweils viermal die höchsten Werte für κ^* . Die Bildung der κ^* -Mittelwerte für die einzelnen Versuchsreihen offenbart, dass V1b mit $\kappa_{V1b}^* = 0,84$ im Mittel die besten Resultate liefert, dicht gefolgt von V2 mit $\kappa_{V2}^* = 0,83$. Aufgrund der schlechten Erkennenleistung resultieren aus V1a mit $\kappa_{V1a}^* = 0,73$ die schlechtesten Werte für den Aufgabenerfolg.

Die Erfolgsmaße auf Nutzerseite E_N (siehe Abbildung 7.21) liefern tendenziell die besseren Ergebnisse für V1, V1b übertrifft V2 in sechs Fällen deutlich. Die durchschnittlichen Werte für das Erfolgsmaß E_N betragen $E_{N_{V1a}} = 0,62$, $E_{N_{V1b}} =$

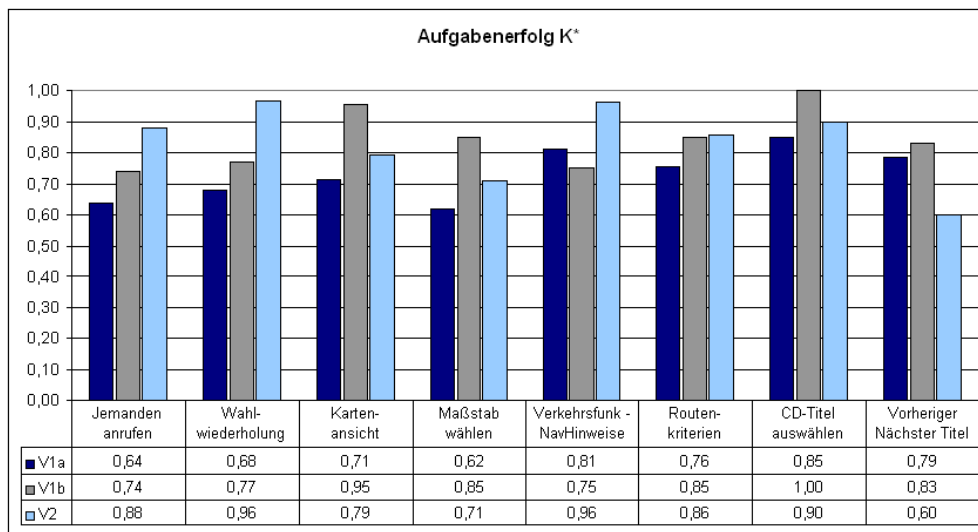
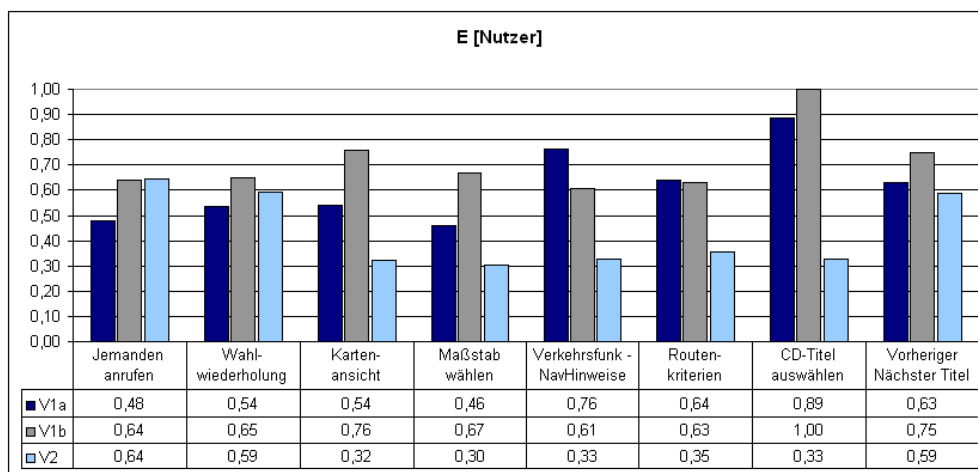
Abbildung 7.20: Aufgabenerfolg κ^* 

Abbildung 7.21: Erfolgsmaß (Nutzer)

0,71 und $E_{N_{V2}} = 0,43$. Bei V1 und insbesondere bei V1b ist der Anteil der den Dialog positiv beeinflussenden Sprecherbeiträge höher als bei V2. Die sehr niedrigen Werte von E_N bei den Aufgaben im Bereich Navigation (Kartenansicht, Maßstab wählen, Navigationshinweise und Routenkriterien) deuten darauf hin, dass die Probanden bei V2 größere Probleme hatten, diese Aufgaben zu lösen, als bei V1.

Die Erfolgsmaße auf Systemseite E_S beschreiben die positiven Beiträge des SDS zu den Dialogen und kennzeichnen erneut die schlechte Spracherkennerleistung bei

V1a (siehe Abbildung 7.22). Mit einem durchschnittlichen Wert von $E_{SV1a} = 0,64$ erhält der Prototyp mit schlechter Erkennerleistung die schlechtesten Werte. V1b und V2 scheiden mit $E_{SV1b} = 0,81$ und $E_{SV2} = 0,82$ nahezu identisch ab, wobei V2 geringfügig bessere Ergebnisse aufweist.

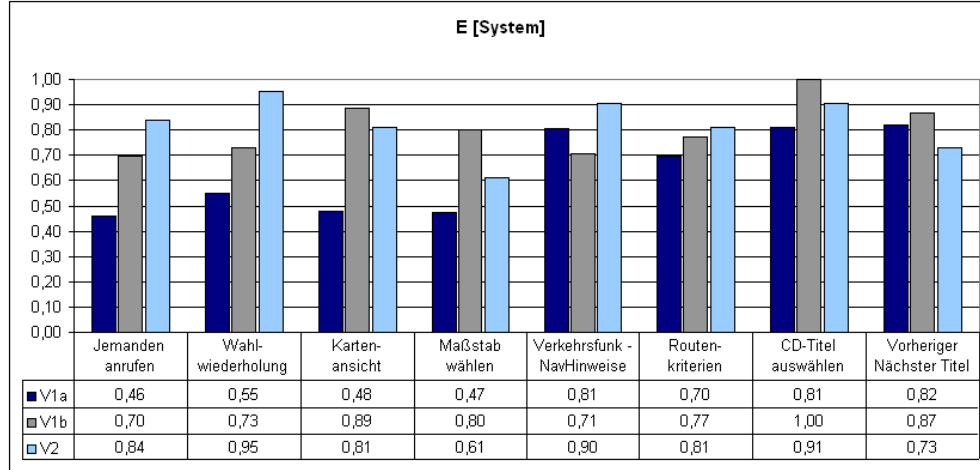


Abbildung 7.22: Erfolgsmaß (System)

Die aus den Versuchsreihen ermittelten Werte können dazu verwendet werden, eine Performanzfunktion zur Vorhersage der Nutzerzufriedenheit zu erstellen. Die besten Gleichungen bestehen für den Prototypen aus folgenden Komponenten:

$$\begin{aligned} \text{US1} = & 0,59 \cdot \text{Gesamtbeiträge} + 0,604 \cdot P(A) - 0,086 \cdot \text{Benutzerbeiträge} - \\ & - 0,088 \cdot \text{Barge-Ins} - 0,002 \cdot \text{Zeit bereinigt [s]} \end{aligned} \quad (7.5)$$

$$\begin{aligned} \text{US2} = & 0,074 \cdot \text{Gesamtbeiträge} + 2,993 \cdot \kappa^* - 3,735 \cdot E_N - \\ & - 0,049 \cdot \text{Benutzerbeiträge} - 0,175 \cdot \text{Optionen} - \\ & - 0,002 \cdot \text{Zeit bereinigt [s]} \end{aligned} \quad (7.6)$$

Da die Performanzfunktion für den Prototyp nur ca. 60% der beobachteten Werte abdeckt ($R_{US1}^2 = 0,585$, $R_{US2}^2 = 0,631$), ist diese nicht zur Vorhersage der Nutzerzufriedenheit geeignet. Eine ausreichend hohe Korrelation zwischen Nutzerzufriedenheit und Performanz des Systems kann nicht verzeichnet werden.

Das Referenzsystem liefert folgende optimale Performanzfunktionen:

$$\begin{aligned} \text{US1} = & -5,529 \cdot E_N + 29,929 \cdot E_S - 0,005 \cdot \text{Gesamtbeiträge} - \\ & - 36,126 \cdot \kappa^* \end{aligned} \quad (7.7)$$

$$\begin{aligned} \text{US2} = & 0,136 \cdot \text{Gesamtbeiträge} + 7,591 \cdot P(A) - \\ & - 0,276 \cdot \text{Optionen} \end{aligned} \quad (7.8)$$

Mit Deckungsgraden von 95% ($R_{US1}^2 = 0,950$) und 80% ($R_{US2}^2 = 0,801$) können die Performanzfunktionen für das Referenzsystem herangezogen werden, jedoch ist auch hier keine ausreichende Signifikanz festzustellen, die eine ausreichend hohe Korrelation zwischen Nutzerzufriedenheit und Performanz nachweist. Bei der Verwendung der identischen Performanzfunktionen für V1 und V2 sinkt der Deckungsgrad R^2 so stark ab, dass die Ergebnisse keine Rückschlüsse mehr auf die Nutzerzufriedenheit zulassen. Die Eignung des PARADISE Evaluierungsframeworks scheint bezüglich der hier vorgestellten Domäne somit fraglich.

7.4.5 Messung mentaler Beanspruchung mittels NASA TLX


Ein wichtiger Aspekt bei der Beurteilung des Prototyps ist die Bewertung der von ihm ausgehenden mentalen Belastung. Die Primäraufgabe im Fahrzeug muss die Fahrzeugbedienung sein, jegliche Sekundäraufgaben müssen hinsichtlich ihrer mentalen Belastung für den Fahrer minimiert werden. Die Bestimmung der subjektiven Arbeitsbelastung und -beanspruchung kann durch verschiedene Messverfahren erreicht werden, z.B. durch die Subjective Workload Assessment Technique (SWAT), den Subjective Mental Effort Questionnaire (SMEQ) oder den NASA Taskload Index (NASA-TLX). Nach Moise (2003) weisen die Ergebnisse des NASA-TLX jedoch die größte Validität und Reliabilität auf und korrelieren stärker mit dem Aufgabenerfolg der Versuchspersonen. Aus diesem Grund erfolgt die Messung der mentalen Beanspruchung mit Hilfe des NASA-TLX. Die Erhebung der Werte erfolgt zunächst durch die Bewertung der Aspekte Mental Demand (Geistige Beanspruchung), Physical Demand (Körperliche Beanspruchung), Temporal Demand (Zeitliche Beanspruchung), Performance (Aufgabenerfüllung), Effort (Anstrengung) und Frustration (Unzufriedenheit) auf einer 18-stufigen Skala (siehe Abbildung 7.23). Nach Scheufler (2002) beziehen sich die ersten drei Dimensionen auf die Anforderungen, die an die Probanden gestellt werden, die letzten drei Dimensionen beschreiben die Interaktionen der Probanden mit den Aufgaben. Der zweite Teil der Datenerfassung besteht aus Paarvergleichen, bei denen 15 Paare aus den o.g. Dimensionen gebildet und gewichtet werden. Pro Paar müssen die Probanden entscheiden, welcher Aspekt stärkeren Einfluss auf das subjektive Befinden hatte.

Von besonderem Interesse sind neben dem durchschnittlichen Workload (Durchschnitt aller Dimensionen) v.a. die Aspekte *Geistige Beanspruchung*, *Aufgabenerfüllung*, *Anstrengung* und *Unzufriedenheit*. Die Dimensionen *Körperliche Beanspruchung* und *Zeitliche Beanspruchung* sind von untergeordnetem Interesse, da während der Versuche keine künstliche Stresssituation herbeigeführt wurde. Abbildung 7.24 zeigt die Ergebnisse der Untersuchung. Der gesamthafte Workload ist für V1b am niedrigsten. Trotz der schlechten Erkennenleistung schneidet V1a

Task Questionnaire - Part 1


Click on each scale at the point that best indicates your experience of the task

Mental Demand




Low High

Physical Demand



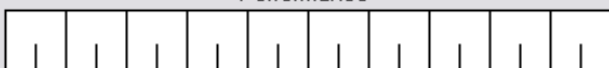
Low High

Temporal Demand



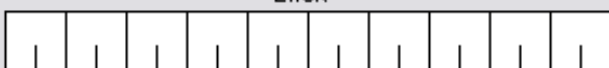
Low High

Performance



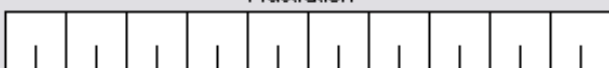
Good Poor

Effort



Low High

Frustration



Low High

Cancel
Continue

Abbildung 7.23: Bewertungsskala des NASA-TLX

noch geringfügig besser ab als V2, der Unterschied ist jedoch marginal. Ein ähnliches Ergebnis liefert die Betrachtung der Geistigen Beanspruchung und der Aufgabenerfüllung. Auch hier liefert V1b die besten Resultate gefolgt von V2 und V1a. Die Dimensionen Anstrengung und Unzufriedenheit sind bei V1b am höchsten, was überraschend ist, da aufgrund der schlechten Erkennerleistung die Unzufriedenheit bei V1a höher zu erwarten gewesen wäre. V2 liefert diesbezüglich die besten Resultate. Mögliche Ursachen können der Einsatz adaptiver oder multimodaler Dialoge sein. Unter Umständen ist der Umgang mit einem komplexeren Bediensystem trotz der positiven Grundbewertung anstrengender als mit dem

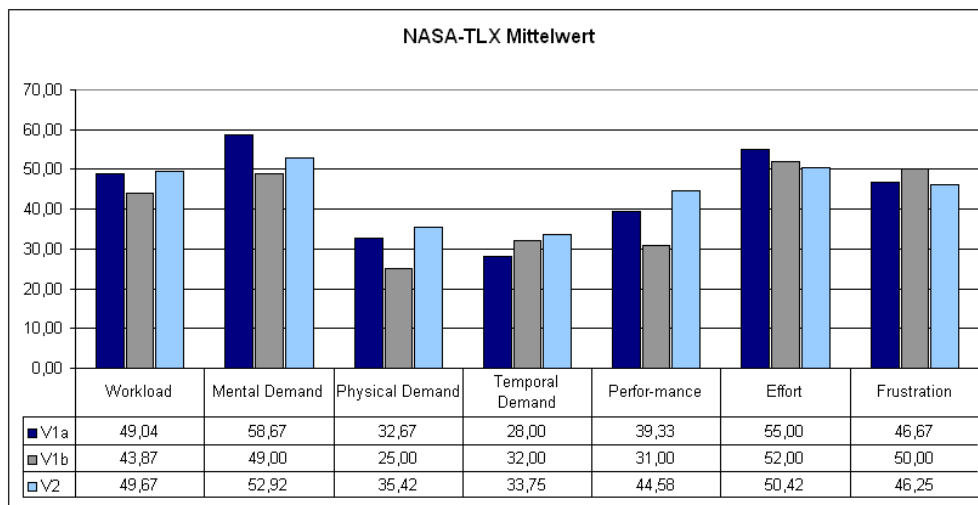


Abbildung 7.24: Mentale Belastung basierend auf dem NASA-TLX

einfacheren iDrive Bediensystem. Da die Konzepte der Adaption nur einem geringen Prozentsatz der Probanden auffielen, könnte dies ein Indiz dafür sein, dass die multimodale Bedienung einen größeren Beitrag zu den Dimensionen Anstrengung und Unzufriedenheit leistet. Dies könnte auch durch die gesteigerte Zahl an Displayblicken pro Minute gestützt werden, die beim Prototyp etwa doppelt so hoch war wie beim iDrive System. Eine genaue Betrachtung dieser Aspekte muss jedoch in umfangreicheren Tests erfolgen.

7.5 Diskussion der Ergebnisse

Die Evaluierung des Prototyps besitzt einen formativen Charakter, als Ziel steht somit im Vordergrund, potenzielle Schwachstellen des Hilfesystems zu identifizieren. Eine umfassende summative Evaluierung erschien nicht zielführend, da aufgrund des geringen Stichprobenumfangs keine statistisch signifikanten Ergebnisse erwartet werden konnten. Die vorliegenden Ergebnisse weisen jedoch Tendenzen auf und tragen dazu bei, Potenziale und Schwachstellen des Hilfesystems zu identifizieren und die aufgestellten Arbeitshypothesen zu untermauern oder zu widerlegen.

1. **Das kontextspezifische Hilfesystem basierend auf adaptiven Komponenten und visueller Unterstützung führt im Gegensatz zum iDrive Hilfesystem zu einer höheren Benutzerzufriedenheit.**

Die Nutzerzufriedenheit US, US1 und US2 war bei Verwendung des Prototypen jeweils höher als beim iDrive Referenzsystem (siehe Abbildung 7.19).

Des weiteren traten beim Prototyp weniger Unklarheiten nach der Verwendung des Hilfesystems auf als beim Referenzsystem (siehe Abbildung 7.12).

2. Die Adaption erhöht die Effizienz beim Auffinden gesuchter Informationen im Hilfesystem.

Hinsichtlich der Aufgabendauer schneidet der Prototyp mit guter Erkennungsleistung stets besser ab als das iDrive System. Interessant ist vor allem, dass die Bearbeitungsdauer bei beiden Hilfesystemen zunächst relativ hoch ist (siehe Abbildung 7.6). Gegen Ende der Tests bleibt die Bearbeitungsdauer beim iDrive System auf einem höheren Niveau, während diese beim Prototyp in höherem Maße absinkt. Dies könnte auf die Adaption bezüglich der Erfahrung eines Benutzers zurückzuführen sein.

Die Bewertung des Aufgabenerfolgs fällt für den Prototyp ebenfalls leicht besser aus (siehe Abbildung 7.20). Neben dem besser Erfolgsmaß E_N (siehe Abbildung 7.21) lassen die besseren κ^* -Werte ebenfalls auf eine Überlegenheit des Prototyps schließen.

3. Eine benutzeradaptive Anpassung der Reihenfolge der Sprachoptionen fällt beiegeeigneter Parameterwahl nicht negativ auf.

Die Auswertung der Fragebögen und Kommentare der Probanden liefern kein Indiz dafür, dass die Adaption der Hilfe zu Irritationen führte (siehe Abbildung 7.10). Die Anpassung der Hilfe war offenbar konservativ genug ausgelegt, um nicht aufzufallen. Auch die selteneren Unklarheiten nach Verwendung der Hilfe lassen den Schluß zu, dass die auftretende Inkonsistenz innerhalb der Dialoge keine negativen Auswirkungen auf die Qualität der Hilfe hat (siehe Abbildung 7.12).

4. Die Erhöhung des Informationsgehalts der Sprachoptionen bei Experten wirkt sich positiv auf die Bedienung aus.

Die Aufgabendauer bei der Bedienung v.a. langer Optionenausgaben (z.B. Routenkriterien) fällt beim Prototyp deutlich geringer aus als beim Referenzsystem (siehe Abbildung 7.6). Dies kann auf die Veränderung der Hilfeausgaben zurückgeführt werden, da hier oftmals bereits eine Anpassung der Länge der Hilfeausgaben eintrat (Benutzer = Experte). Zudem wanderten bekannte Sprachbefehle in nachgelagerte Hilfeebenen, so dass die gesuchten Kommandos häufig schon zu Beginn der Hilfe zu finden waren. Beim iDrive-System mussten je nach Position des Kommandos im Hilfetext die gesamten Sprachausgaben angehört werden. Des weiteren entfielen beim Übergang eines Benutzers vom Anfänger zum Experten die Beispiele in den Sprachausgaben, was sich ebenfalls positiv auf die Aufgabendauer auswirkte, den Aufgabenerfolg jedoch nicht negativ beeinflusste (siehe Abbildung 7.20 und 7.21).

5. **Die mehrstufig gegliederte Ausgabe von akustischer Information reduziert die mentale Belastung im Gegensatz zu einer ungegliederten Ausgabe.**

Die geistige Beanspruchung war beim Prototyp mit gutem Erkennen geringer als beim Referenzsystem (siehe Abbildung 7.24). Auch für die gesamthafte Belastung der Probanden lassen sich ähnliche Ergebnisse nachweisen. Die bei komplexeren Aufgaben kürzeren Aufgabendauern (siehe These 4) stützen diese These ebenfalls.

6. **Multimodale Interaktion führt zu effizienteren Dialogen hinsichtlich der Bediengeschwindigkeit sowie zu höherer Benutzerzufriedenheit.**

Die Probanden verwendeten die visuelle Hilfe v.a. in Situationen, in denen die Beanspruchung durch die Primäraufgabe sie von der akustischen Hilfe ablenkte. Auch die explizite Hervorhebung der vorgelesenen Sprachbefehle vermittelte den Probanden ein positives Gefühl für die Dialogführung und bestärkte den positiven Gesamteindruck der visuellen Hilfe (siehe Abbildung 7.13). Obwohl multimodale Interaktionen nur in sehr seltenen Fällen durchgeführt wurden, beurteilten die Probanden diese Interaktionsmöglichkeit überwiegend positiv (siehe Abbildung 7.14).

Bei der Analyse der Blickzuwendungen weisen die beiden Systeme nur geringe Unterschiede hinsichtlich der Verweildauern von mehr als zwei Sekunden auf CID oder Tacho auf, der Prototyp liefert geringfügig bessere Werte. Bei der Verteilung der Blickzuwendungen erzielten beide Systeme ebenfalls ähnliche Werte (ca. 30%), wobei das iDrive-System aufgrund der fehlenden grafischen Unterstützung jedoch die geringfügig besseren Werte lieferte (siehe Abbildung 7.9).

7. **Die zusätzliche visuelle Ausgabe der Hilfe forciert die Umstellung von Sprachbedienung auf haptische Bedienung.**

Der Anteil aller haptischen Interaktionen bezogen auf die Gesamtmenge an Interaktionen beträgt lediglich 2%. Dieser sehr geringe Anteil der Controllerverwendung ist zudem gleichmäßig über die einzelnen Versuche verteilt, eine Zunahme der Controllerverwendung mit fortschreitender Erfahrung in der Systembedienung ist nicht zu verzeichnen. Auch die Verwendung der Push-In Funktionalität blieb hinter den Erwartungen zurück, da ein Push-In lediglich als robuste Eingabemethode im Falle einer schlechten Spracherkennung verwendet wurde, nicht jedoch in unproblematischen Dialogen.

8. **Kontextunabhängige Hilfe wird in unbekannten Domänen gegenüber kontextabhängiger Hilfe bevorzugt.**

Die kontextunabhängige Hilfe wurde von den Probanden positiv bewertet (siehe Abbildung 7.15). In den Szenarien, in denen die kontextunabhängige

Hilfe eingesetzt wurde, erreichten alle Probanden das gewünschte Ziel. Eine vermehrte Verwendung dieser Form der Hilfe konnte jedoch nicht festgestellt werden. Das Kommando „Optionen“ wurde überwiegend als erstes verwendet, um weitergehende Informationen zu erhalten. Die Hemmschwelle, mit dem SDS über natürlichsprachliche Eingaben und Fragestellungen zu kommunizieren, war bei den Probanden offenbar relativ hoch. Vor allem in den Fällen, in denen der erste Versuch einer natürlichsprachlichen Hilfeanfrage scheiterte, bevorzugten die Probanden die robustere Methode der Optionen. Die geringere Verwendungshäufigkeit könnte auch auf die technischen Rahmenbedingungen zurückzuführen sein, die eine robuste Sprachverarbeitung nicht immer gewährleisten konnten.

7.6 Fazit

Die Evaluierung des prototypischen Hilfesystems erfolgte im direkten Vergleich zum Hilfesystem des iDrive SDS. Da der Prototyp des Hilfesystems auf dem Anzeige-Bedienkonzept der kommenden Baureihen basiert, unterlag der Prototyp zum Zeitpunkt der Tests der Geheimhaltung. Die Untersuchung des Prototyps fand daher in einer isolierten Testumgebung im Fahrsimulator statt, die Tests des iDrive-Systems wurden in einem Serienfahrzeug durchgeführt. Die Versuche wurden im Between-Subject-Design geplant, d.h. die beiden Hilfesysteme wurden unabhängig voneinander von zwei Versuchsgruppen beurteilt. Die Probanden mussten jedoch die gleichen Aufgaben bearbeiten. Bei der Auswertung der Ergebnisse wurde die Versuchsgruppe des Prototyps in zwei Untergruppen aufgeteilt, da die anfänglich schlechte Erkennerleistung eines neuen Spracherkenners die Beurteilung der Probanden negativ beeinflusste.

Die Analyse der Videoaufzeichnungen und Fragebögen lieferte ein überwiegend positives Resultat zu Gunsten des Prototyps. Sowohl die beobachteten Systeminteraktionen als auch die subjektiven Bewertungen durch die Probanden heben das prototypische Hilfesystem positiv vom iDrive-System ab. Die Anwendung des PARADISE Evaluierungsframeworks spiegelt die positiven Resultate ebenfalls wider. Die Nutzerzufriedenheit ist beim Prototyp tendenziell stets höher als beim iDrive-System. Gleiches gilt für das Erfolgsmaß seitens der Benutzer (wieviel hat der Benutzer zum Erreichen eines Dialogziels beigetragen) und den Aufgabenerfolg κ^* (für die Versuchsreihe mit vergleichbarer Erkennerleistung zum iDrive-System). Lediglich das Erfolgsmaß auf Systemseite liefert minimal bessere Resultate für das iDrive-System. Die Bildung einer einheitlichen Performanzfunktion für den direkten Vergleich beider Systeme liefert jedoch ein unbefriedigendes Resultat. Die Eignung des PARADISE Evaluierungsframeworks für die Bewertung von SDS in automotiven Umgebungen scheint somit fraglich.

Die Messung der durchschnittlichen mentalen Belastung mittels NASA-TLX liefert für den Prototyp mit gutem Spracherkenner die besten Ergebnisse, gefolgt vom iDrive-System und dem Prototyp mit schlechter Erkennenleistung. Ein ähnliches Resultat ergibt sich für die Dimensionen der geistigen Beanspruchung und der Aufgabenerfüllung. Die mangelhafte Erkennenleistung wirkt sich somit nachweislich besonders stark auf die mentale Belastung der Benutzer aus.

Kapitel 8

Zusammenfassung und Ausblick

Die Einführung von SDS im Fahrzeug markiert einen wichtigen Punkt in der Entwicklung benutzerfreundlicher Bedienkonzepte. Die Nachteile klassischer visuell-haptischer Anzeige-Bedienkonzepte (Ablenkung vom Verkehrsgeschehen, komplexe Menüstrukturen) werden durch den Einsatz von SDS z.T. gelöst, jedoch treten neue Herausforderungen beim Entwurf von SDS auf. Der Umfang an Funktionen, die im Fahrzeug zur Verfügung stehen, nimmt stetig zu. Neben klassischen Geräten wie dem Navigationssystem, dem Telefon, Radio oder CD Player werden stets neue Funktionen integriert, z.B. Digitales Radio, Online-Dienste oder weitere Fahrerassistenzsysteme. Der Erfolg der SDS hängt hierbei entscheidend vom Verhalten des Systems ab, insbesondere beim Erstkontakt. Die Erfahrung zeigt, dass die Benutzer sich bei den ersten Verwendungen eines SDS ein Urteil über das System bilden. Verhält sich das SDS nicht erwartungskonform oder bietet es keine verständliche Hilfe an, sinkt die Akzeptanz gegenüber dem SDS. Die Integration eines wirkungsvollen Hilfesystems leistet somit einen entscheidenden Beitrag zum Erfolg eines SDS. Im Zuge des steigenden Funktionsumfangs stoßen die bisherigen Hilfesysteme jedoch an ihre Grenzen, weshalb die Einführung neuer Konzepte notwendig ist.

Kapitel 8.1 beschreibt, inwiefern das Wissen des Benutzers modelliert werden kann und wie dies dazu beitragen kann, die Hilfe zu individualisieren und an die Bedürfnisse des Benutzers anzupassen. Die Berücksichtigung des Erfahrungsgrades des Benutzers im Hinblick auf die allgemeine Systembedienung liefert weiteres Potenzial zur Verbesserung des Hilfesystems. Die Anpassung der Dialogstrategie des Hilfesystems anhand der Systemerfahrung wird ebenfalls in Kapitel 8.1 thematisiert. In Kapitel 8.2 werden die multimodalen Komponenten des Hilfesystems aufgeführt, die dazu beitragen, einen intuitiveren und natürlicheren Dialogfluss zu gewährleisten. Die kontextunabhängige Hilfe wird in Kapitel 8.3 erläutert. Abschließend gibt Kapitel 8.4 einen Überblick über die Evaluierung des prototypischen Hilfesystems und die Ergebnisse des Versuche. Fragestellungen, die in Bezug auf die Entwicklung von Hilfesystemen für SDS offen blieben oder neu hinzukamen werden abschließend in Kapitel 8.5 diskutiert.

8.1 Anpassung der Hilfeinhalte und der Dialogstrategie an das Wissen und den Erfahrungsgrad des Benutzers

Die Neukonzeption eines Hilfesystems für SDS im Fahrzeug unterliegt den Forderungen nach kürzeren Dialogen, einer gesteigerten Nutzerfreundlichkeit und einer sinkenden mentalen Belastung. Die teilweise langen Hilfedialoge des iDrive SDS wurden zunächst in mehrere Teildialoge basierend auf verschiedenen Hilfeebenen untergliedert. Dies wirkt sich einerseits positiv auf die Behaltensleistung aus (Serialer Positionseffekt) und trägt andererseits direkt zur Verkürzung der Dialoge bei. Die Inhalte der kontextspezifischen Hilfe werden an das Wissen des Benutzers angepasst. Mit Hilfe des Potenzgesetzes der Übung und einer Vergessenskurve wird dieses Wissen mathematisch erfasst und in ein Benutzermodell überführt. Der Lernerfolg eines Benutzers hängt nach dem Potenzgesetz der Übung primär von der Zeit bei der ersten Verwendung einer Funktion ab. Durch die Bestimmung dieser Zeit kann errechnet werden, wie oft eine bestimmte Funktion wiederholt werden muss, um als gelernt zu gelten. Zu diesem Zweck erhält jede Funktion einen Index, der die Häufigkeit der Verwendung wiedergibt. Wurde eine Funktion häufiger als der berechnete Schwellenwert verwendet, gilt die Funktion als gelernt und wandert an das Ende der Hilfeausgabe. Wird die entsprechende Funktion jedoch über einen längeren Zeitraum nicht mehr verwendet, muss davon ausgegangen werden, dass die Funktion u.U. vergessen wurde. Diesem Vergessenprozess wird durch die Anwendung einer Vergessenskurve auf den Index einer Funktion Rechnung getragen. Fällt der Index im Laufe der Zeit wieder unter den zugehörigen Schwellenwert, wird die Funktion wieder an ihre ursprünglichen Position in der Hilfeausgabe verschoben. Aufbauend auf diesem Benutzermodell beinhaltet die Hilfe primär Informationen über diejenigen Funktionen, die ein Benutzer nicht so gut kennt oder selten verwendet. Häufig verwendete Funktionen und Funktionen, die bereits bekannt sind und als gelernt betrachtet werden, befinden sich am Ende der Hilfeausgabe.

Die Länge der einzelnen Teildialoge der Hilfe leitet sich aus dem Erfahrungsgrad des Benutzers im Umgang mit dem SDS ab. Zur Bestimmung des Erfahrungsgrades dient ein Verfahren, dessen Grundprinzip auf der Arbeit von Hassel (2006) beruht. Dabei werden Parameter des Umgangs eines Benutzers mit dem System erfasst (Hilfe- und Optionsanfragen, Timeouts, Erkennenfehler und Barge-Ins). Aus diesen Werten und einem Gewichtsvektor \vec{UM}_G kann durch Bildung des Skalarprodukts ein generisches Benutzermodell \vec{UM} (Anfänger, Experte) bestimmt werden. Anfänger erhalten kürzere Hilfedialoge bestehend aus vier Informationseinheiten pro Dialogschritt, Expertendialoge hingegen enthalten fünf Informationseinheiten. Zudem entfallen bei Experten die Beispiele in den Hilfeausgaben.

8.2 Multimodale Hilfefunktionalität

Die Einführung eines SDS beseitigt zwar z.T. bestehende Nachteile rein visuell-haptischer Bedienkonzepte, jedoch kann der Aufruf einer Funktion per Sprache je nach Dialogziel u.U. umständlicher sein (z.B. Blättern in langen Listen, kurze Bestätigung per Druck auf den Controller, etc.). Ziel muss es also sein, für verschiedene Dialogziele die optimale Ein- und Ausgabemodalität zur Verfügung zu stellen. Dies kann entweder isoliert geschehen (exklusiv-multimodale Interaktion) oder verknüpft (alternierend-multimodale bzw. simultan-multimodale Interaktion). Das prototypische Hilfesystem stellt dem Benutzer exklusiv-multimodale und alternierend-multimodale Interaktionsvarianten zur Verfügung. Je nach Präferenz kann ein per Sprache begonnener Dialog mit dem Controller fortgeführt oder beendet werden. Die akustischen Sprachausgaben werden durch eine visuelle Repräsentation ergänzt. Die visuelle Ausgabe erfolgt zeitlich synchron zur Sprachausgabe, was während der Hilfe eine Hervorhebung des aktuell vorgelesenen Sprachkommandos ermöglicht. In Analogie zum Barge-In hat der Benutzer die Möglichkeit, einen Push-In durchzuführen und eine vorgelesene Funktion per Controller aufzurufen.

8.3 Kontextunabhängige Hilfefunktionalität

Die kontextunabhängige Hilfe ist speziell für den Erstkontakt mit dem gesamten SDS oder unbekannten Teilbereichen konzipiert. Die Benutzer haben die Möglichkeit, natürlichsprachliche Systemanfragen zu formulieren (z.B. „Wie kann ich Radio hören?“), worauf das Hilfesystem eine explizite Anleitung ausgibt. Dazu wird die Struktur des SDS mit Hilfe eines ungerichteten zyklischen Graphen abgebildet. Beim Aufruf der Hilfe wird der aktuelle Kontext des Benutzers im System ermittelt und mit Hilfe des Algorithmus von Dijkstra der kürzeste Weg zum Dialogziel errechnet. Das Hilfesystem ermittelt die notwendigen Kommandos, die der Benutzer sprechen muss, um sein Dialogziel zu erreichen.

8.4 Evaluierung des Hilfesystems

Das prototypische Hilfesystem wurde mit dem Hilfesystem des iDrive Anzeige-Bedienkonzepts verglichen. Die Tests wurden mit zwei unabhängigen Testgruppen bestehend aus 15 bzw. zwölf Versuchspersonen durchgeführt. Die Untersuchung des prototypischen Hilfesystems fand aus technischen und organisatorischen im Fahrsimulator statt, das iDrive-System wurde in einem Serienfahrzeug getestet. Alle Probanden erhielten jedoch die gleiche Aufgabenstellung.

Die Versuche wurden mittels Videokameras aufgezeichnet, um eine nachträgliche Auswertung verschiedener Interaktionsparameter zu ermöglichen. Die Analyse der Videoprotokolle und der Fragebögen lieferten ein überwiegend positives Resultat zu Gunsten des Prototyps. Sowohl die beobachteten Systeminteraktionen als auch die subjektiven Bewertungen durch die Probanden heben das prototypische Hilfesystem positiv vom iDrive-System ab. Hinsichtlich der Aufgabendauer, der Anzahl der Hilfeaufrufe, der Länge der Hilfeausgaben und der auftretenden Unklarheiten nach dem Aufruf der Hilfe weist der Prototyp bessere Resultate auf als das iDrive-System. Die Messung der gesamthaften mentalen Belastung mittels NASA-TLX liefert für den Prototyp mit gutem Spracherkenner die besten Ergebnisse. Ein ähnliches Resultat ergibt sich für die Dimensionen der geistigen Beanspruchung und der Aufgabenerfüllung. Die Untersuchung der Gebrauchstauglichkeit des Prototyps fand mit Hilfe des PARADISE Evaluierungsframeworks statt. Die Nutzerzufriedenheit ist beim Prototyp tendenziell stets höher als beim iDrive-System. Gleiches gilt für das Erfolgsmaß seitens der Benutzer (wieviel hat der Benutzer zum Erreichen eines Dialogziels beigetragen) und bei gleichen Systemvoraussetzungen für den Aufgabenerfolg κ^* . Lediglich das Erfolgsmaß auf Systemseite liefert minimal bessere Resultate für das iDrive-System. Die Bildung einer einheitlichen Performanzfunktion für den direkten Vergleich beider Systeme liefert jedoch ein unbefriedigendes Resultat.

8.5 Ausblick

Die in dieser Arbeit vorgestellten Konzepte zur Anpassung der kontextspezifischen Hilfe an das Wissen und die Erfahrungen von Benutzern basieren auf der Zuordnung von individuellen Benutzerprofilen. Auf welche Weise dies geschehen ist nicht Bestandteil dieser Arbeit. Denkbar wäre die Speicherung des persönlichen Profils auf einer Keycard, wie sie z.T. schon in Fahrzeugen zum Einsatz kommt. Dies birgt jedoch die Problematik der eindeutigen Zurodnung des Profils, z.B. falls das Fahrzeug von mehreren Personen in einem Haushalt oder Fuhrpark verwendet wird. Günstiger erscheint die Anbindung an biometrische Identifikations- und Authentifizierungsverfahren. In Heckner (2005) wird die Verwendung der videobasierten Personenidentifikation (Gesichtserkennung) als Möglichkeit zur Individualisierung und Authentifizierung innerhalb von Mensch-Maschine-Schnittstellen erläutert. Dieses Verfahren wäre sehr gut zur Profilzuordnung geeignet, da seitens des Benutzers keinerlei Eingaben oder Interaktionen notwendig sind.

Die Anpassung der Hilfeausgaben berücksichtigt bisher lediglich kognitionspsychologische Aspekte (Lernen und Vergessen) und individuelle Fähigkeiten im Umgang mit dem SDS. Sinnvoll wäre jedoch auch die Adaption hinsichtlich der körperlichen und geistigen Belastung im Moment der Dialogverwendung. Ist der Fahrer z.B. einer Stresssituation ausgesetzt, ist der Fahrer genervt, erfordert der Stra-

ßenverkehr höchste Konzentration oder sind Anzeichen von Übermüdung erkennbar? Jeder dieser Aspekte spielt sowohl bei der Fahrzeugführung als auch bei der Bedienung eines Mensch-Maschine-Systems eine wichtige Rolle. Die Berücksichtigung der o.g. Aspekte könnte durch die Erfassung biometrischer Daten erfolgen, z.B. durch die Messung der Hautleitfähigkeit oder der Herzfrequenz, was Rückschlüsse auf die Stresssituation zulassen könnte. Die Auswertung sensorischer Daten (z.B. von Radarsensoren, GPS) oder von Verkehrsinformationen könnte dazu beitragen, die Verkehrssituation und -dichte zu bewerten. Basierend auf diesen Informationen könnten Dialoge entsprechend angepasst werden, um die Belastung des Fahrers zu minimieren.

Großes Entwicklungspotenzial bietet die kontextunabhängige Hilfe. Mit weiter verbesserten Spracherkennern ist es möglich, die Verarbeitung natürlich- oder spontansprachlicher Äußerungen möglich zu machen. Insofern könnten Hilfeanfragen wesentlich freier erfolgen und wären an kein festes Schema gebunden. Natürlichsprachliche Dialoge bieten dem Benutzer ein hohes Maß an Freiheit hinsichtlich der Dialogführung, Hilfesysteme für die Form der Interaktion müssen jedoch dynamischer gestaltet werden, um das individuelle Informationsbedürfnis zu befriedigen. Denkbar wären framebasierte Ansätze bei der Entwicklung von Hilfesystemen, d.h. das Hilfesystem ermittelt die vom Benutzer eingegebenen Informationen und bewertet mögliche Dialogziele sowie die dazu noch fehlenden Informationen. Hilfesysteme müssten somit enger mit dem regulären SDS verzahnt werden.

Im Rahmen der kontextunabhängigen Hilfe wäre zudem die Anpassung der angebotenen Informationsdichte denkbar. Je nach Erfahrungsgrad oder Belastungssituation sollte die Anleitung für ein bestimmtes Dialogziel detaillierter oder komplexer ausfallen. Anfänger benötigen u.U. eine Schritt-für-Schritt-Anleitung, um das gewünschte Dialogziel zu erreichen. Experten, die mit dem SDS oder bestimmten Teilbereichen des Systems sehr gut vertraut sind, benötigen hingegen keine ausführliche Anleitung. Kommandos, die sie bereits kennen, könnten in der Hilfeausgabe zusammengefasst werden, ausführlichere Anleitungen könnten sich auf Themengebiete beschränken, die sie nicht gut kennen. Denkbar wäre in diesem Zusammenhang auch eine visuelle Unterstützung durch Animationen oder Screenshots. Hierbei müsste jedoch untersucht werden, inwieweit sich dies auf die mentale Belastung und die Ablenkung vom Straßenverkehr auswirkt.

Ein wichtiger Aspekt bei der Einführung eines adaptiven Systems sind Langzeitversuche, um Erkenntnisse über die Akzeptanz und die Effektivität eines solchen Systems zu sammeln und im Rahmen einer summativen Evaluierung zu verifizieren. Diese Art der Information kann nicht in einer kontrollierten Laborumgebung oder in kurzen Erprobungstests erhoben werden. Um aussagekräftige Resultate zu erhalten, sind Versuche unter realen Bedingungen mit einer größeren Anzahl von Versuchspersonen über einen längeren Zeitraum nötig. Nur so können die Schachstellen eines Systems aufgedeckt und behoben werden.

Literaturverzeichnis

- [Althoff et.al. 2001] ALTHOFF, Frank ; MCGLAUN, Gregor ; SPAHN, Gunter ; LANG, Manfred: Combining Multiple Input Modalities for Virtual Reality Navigation-A User Study. In: *Proceedings of HCI 2001: 9th International Conference on Human Computer Interaction*. New Jersey : Lawrence Erlbaum Associates, 2001, S. 47–49
- [Anderson 2001] ANDERSON, J.R.: *Kognitive Psychologie*. Heidelberg : Spektrum Akademischer Verlag, 2001
- [Ardissono 1996] ARDISSONO, Liliana: *Dynamic User Modeling and Plan Recognition in Dialogue*, Universita di Torino, Dissertation, 1996
- [Audi 2007] AUDI: *Der Audi Q7 - Innenansicht*. 2007. – URL http://www.audi.de/audi/de/de2/neuwagen/q7/galerie/360_innenansicht.html. – Letzter Zugriff am 25.06.2007
- [Aust et.al. 1995] AUST, Harald ; OERDER, Martin ; SEIDE, Frank ; STEINBISS, Volker: The Philips Automatic Train Timetable Information System. In: *Speech Communication* 17 (1995), Nr. 3-4, S. 249–262
- [Baddeley 1999] BADDELEY, Alan: *Essentials of Human Memory*. Hove, East Sussex : Psychology Press Ltd., 1999
- [Baggen und Hemmerling 2000] BAGGEN, R. ; HEMMERLING, S.: Evaluation von Benutzbarkeit in Mensch-Maschine-Systemen. In: JÜRGENSOHN, T. (Hrsg.) ; KOLREP, H. (Hrsg.): *Mensch-Maschine-Systemtechnik*. Symposion Publishing, 2000, S. 232–284
- [Bauer 1988] BAUER, Joachim: *Konzepte und Prototypen interaktiver Hilfesysteme*, Universität Stuttgart, Dissertation, 1988
- [Bauer et.al. 1993] BAUER, Mathias ; BIUNDO, Susanne ; DENGLER, Dietmar ; KOEHLER, Jana ; PAUL, Gabriele: PHI - A Logic-Based Tool for Intelligent Help Systems. In: BAJCSY, Ruzena (Hrsg.): *Proceedings of the 13th International Joint Conference on Artificial Intelligence*, Morgan Kaufmann, August 1993, S. 460–466
- [Bengler 1995] BENGLER, Klaus: *Gestaltung und experimentelle Untersuchung unterschiedlicher Präsentationsformen von Wegleitungsinformation in Kraftfahrzeugen*, Universität Regensburg, Dissertation, 1995

- [Bengler 2001] BENGLER, Klaus: Aspekte der multimodalen Bedienung und Anzeige im Automobil. In: JÜRGENSOHN, Thomas (Hrsg.) ; TIMPE, Klaus-Peter (Hrsg.): *Kraftfahrzeugführung*. Berlin, Heidelberg, New York, Barcelona, Budapest, Hong Kong, London, Mailand, Paris, Tokyo : Springer Verlag, 2001, S. 195–205
- [Bengler et. al. 2002] BENGLER, Klaus ; HERRLER, Michael ; KÜNZNER, Hermann: Usability Engineering bei der Entwicklung von iDrive. In: *it - Information Technology* 44 (2002), Nr. 3, S. 145–152
- [Benoit 1997] BENOIT, Christian: Audio-Visual and Multimodal Speech Systems. In: GIBBON, Dafydd (Hrsg.): *Handbook of Standards and Resources for Spoken Language Systems*. Mouton de Gruyter, 1997
- [Benyon und Murray 1993] BENYON, David ; MURRAY, Dianne: Applying User Modelling to Human-Computer Interaction Design. In: *AI Review* 43 (1993), Nr. 6, S. 43–69
- [Bernsen et. al. 1998] BERNSEN, Nils O. ; DYBKJÆR, Hans ; DYBKJÆR, Laila: *Designing Interactive Speech Systems: From First Ideas to User Testing*. Berlin, Heidelberg, New York, Barcelona, Budapest, Hong Kong, London, Mailand, Paris, Tokyo : Springer Verlag, 1998
- [Biemans et. al. 2002] BIEMANS, M. ; SETTEN, M. van ; VLIET, H. van ; ALBERINK, M. ; EERTINK, H. ; HEER, J. de ; KRANENBURG, H. van ; KRUSE, H. ; POOT, H. de ; REITSMA, J. ; SLAGTER, R. ; TEEUW, W. ; VEENSTRA, M.: Adaptation (An Integrated View on Adaptation in Telematics) / Telematica Instituut Enschede. 2002 (TI/RS/2002/042). – Forschungsbericht
- [Biundo et. al. 1992] BIUNDO, S ; DENGLER, D ; KÖHLER, S: Deductive Planning and Plan Reuse in a Command Language Environment. In: *Proceedings of the 10th European Conference on Artificial Intelligence*, 1992, S. 628–632
- [BMW Group 2007] BMW GROUP: *BMW Mediapool*. Nicht öffentlich zugänglich. 2007. – URL <https://mediapool.bmwgroup.com>. – Letzter Zugriff am 25.06.2007
- [Bohnenberger 2001] BOHNENBERGER, Thorsten: Dialog Strategies for Adaptive Help Systems. In: *9th GI-Workshop on Adaptivity and User Modeling in Interactive Software Systems*, 2001
- [Brey und Salmen 2003] BREY, Thomas ; SALMEN, Angelika: *Multimodales Interaktionsmanagement*. 2003. – Konzeptstudie für die Robert Bosch GmbH
- [Carstensen et. al. 2004] CARSTENSEN, Kai-Uwe (Hrsg.) ; EBERT, Christian (Hrsg.) ; ENDRISS, Cornelia (Hrsg.) ; JEKAT, Susanne (Hrsg.) ; KLABUNDE, Ralf (Hrsg.) ; LANGER, Hagen (Hrsg.): *Computerlinguistik und Sprachtechnologie*. Heidelberg : Spektrum Akademischer Verlag, 2004

- [Cisco Systems Inc. et.al. 2002] CISCO SYSTEMS INC. ; COMVERSE INC. ; INTEL CORPORATION ; MICROSOFT CORPORATION ; PHILIPS ELECTRONICS N.V. ; SPEECHWORKS INTERNATIONAL INC.: *Speech Application Language Tags (SALT)*. 2002. – URL <http://www.saltforum.org/saltforum/downloads/SALT1.0.pdf>. – Letzter Zugriff: 21.01.2007
- [Cohen et.al. 2000] COHEN, Philip ; MCGEE, David ; CLOW, Josh: The Efficiency of Multimodal Interaction for a Map-based Task. In: *Proceedings of the 6th Applied Natural Language Processing Conference*, 2000, S. 331–338
- [Comité Européen de Normalisation 1996] COMITÉ EUROPÉEN DE NORMALISATION: *Ergonomische Anforderungen für Bürotätigkeiten mit Bildschirmgeräten, Teil10: Grundsätze der Dialoggestaltung (DIN EN 9241-10)*. 1996
- [Davis et.al. 2003] DAVIS, Hasker ; SMALL, Scott ; STERN, Yaakov ; MAYEUX, Richard ; SIMEONFELDSTEIN ; KELLER, Frederick: Acquisition, Recall and Forgetting of Verbal Information in Long-term Memory by Young, Middle-aged, and Elderly Individuals. In: *Cortex* 39 (2003), Nr. 4-5, S. 1063–1091
- [Dix et.al. 2004] DIX, Alan ; FINLAY, Janet ; ABOWD, Gregory ; BEALE, Russell: *Human Computer Interaction*. London : Prentice Hall, 2004
- [Dzida 1983] DZIDA, Wolfgang: Das IFIP-Modell für Benutzerschnittstellen. In: *Office Management Sonderheft* (1983), S. 6–8
- [Ebbinghaus 1885] EBBINGHAUS, Hermann: *Ueber das Gedächtniss*. 1885
- [Erke et.al. 2001] ERKE, H. ; HARMS, D. ; WALLEWEIM, A.: VODIS (Voice Operated Driver Information System) - Das Auto, das aufs Wort gehorcht. In: ZENTRUM FÜR VERKEHR DER TECHNISCHEN UNIVERSITÄT BRAUNSCHWEIG (Hrsg.): *Automatisierungs- und Assistenzsysteme für Transportmittel*. VDI, 2001, S. 112–125
- [Findlater und McGrenere 2004] FINDLATER, Leah ; MCGRENERE, Joanna: A Comparison of Static, Adaptive, and Adaptable Menus. In: *Proceedings of the 2004 Conference on Human Factors in Computing Systems*. New York : ACM Press, 2004, S. 89–96
- [Fischer 2001] FISCHER, Gerhard: User Modeling in Human-Computer Interaction. In: *User Modeling and User-Adapted Interaction* 11 (2001), S. 65–86
- [Fukubayashi et.al. 2006] FUKUBAYASHI, Yuichiro ; KOMATANI, Kazunori ; OGATA, Tetsuya ; OKUNO, Hiroshi: Dynamic Help Generation by Estimating User's Mental Model in Spoken Dialogue Systems. In: *Proceedings of the International Conference on Spoken Language Processing*, 2006, S. 1946–1949
- [Geiser 1990] GEISER, Georg: *Mensch-Maschine-Kommunikation*. München : Oldenbourg Verlag, 1990

- [Glass und Seneff 2003] GLASS, James ; SENEFF, Stephanie: Flexible and Personalizable Mixed-Initiative Dialogue Systems. In: *Proceedings of HLT-NAACL 2003 Workshop on Research Directions in Dialogue Processing*, 2003, S. 19–21
- [Gorrell 2003] GORRELL, Genevieve: *Language Modelling for Spoken Dialogue Systems; Grammar-Based and Robust Approaches Compared and Contrasted*. 2003. – URL http://www.speech.kth.se/~matsb/speech_rec_course_2003/papers/Genevieve_G/speech_course_essay.pdf. – Letzter Zugriff: 21.01.2007
- [Green 1985] GREEN, M.: Report on Dialogue Specification Tools. In: PFAFF, Günther (Hrsg.): *User Interface Management Systems*. Berlin, Heidelberg, New York, Barcelona, Budapest, Hong Kong, London, Mailand, Paris, Tokyo : Springer Verlag, 1985, Kap. Report on Dialogue Specification Tools, S. 9–20
- [Gründl 2006] GRÜNDL, Martin: *Nutzung von iDrive-Funktionen durch BMW 5er Fahrer*. 2006. – Interne Studie
- [Hagen 1999] HAGEN, Eli: An Approach to Mixed Initiative Spoken Information Retrieval Dialogue. In: *User Modeling and User-Adapted Interaction* 9 (1999), Nr. 1, S. 167–213
- [Hagen et. al. 2004] HAGEN, Eli ; SAID, Tarek ; ECKERT, Jochen: *Spracheingabe im neuen BMW 6er*. ATZ. 2004
- [Haller 2003] HALLER, Rudolf: The Display and Control Concept iDrive - Quick Access to All Driving and Comfort Functions. In: *ATZ/MTZ Extra (The New BMW 5-Series)* (2003), S. 51–53
- [Hamberger und Mauter 2003] HAMBERGER, Werner ; MAUTER, G.: Audi Multi Media Interface (MMI): Neue Spezifikationsmethoden zur interdisziplinären Bedienkonzeptentwicklung. In: *VDI-Berichte: Der Fahrer im 21. Jahrhundert (Anforderungen, Anwendungen, Aspekte für Mensch-Maschine-Systeme)* Bd. 1768. Düsseldorf : VDI-Verlag, 2003, S. 217–233
- [Hanrieder und Hamerich 2004] HANRIEDER, Gerhard ; HAMERICH, Stefan: Modeling Generic Dialog Applications for Embedded Systems. In: *Proceedings of the 8th International Conference on Spoken Language Processing*, 2004, S. 237–240
- [Hasebrook 1995] HASEBROOK, Joachim: *Multimedia Psychologie*. Heidelberg Berlin Oxford : Spektrum Akademischer Verlag, 1995
- [Hassel 2006] HASSEL, Liza: *Adaption und Evaluation von Sprachbediensystemen im Automobilbereich*, Ludwig-Maximilians-Universität München, Dissertation, 2006
- [Hastie et. al. 2002] HASTIE, Helen W. ; JOHNSTON, Michael ; EHLEN, Patrick: Context-Sensitive Help for Multimodal Dialogue. In: *Proceedings of the 4th IEEE International Conference on Multimodal Interfaces*, 2002, S. 93–98

- [Heathcote et. al. 2002] HEATHCOTE, Andrew ; BROWN, Scott ; MEWHORT, D. J. K.: The Power Law Repealed: The Case for an Exponential Law of Practice. In: *Psychonomic Bulletin and Review* 7 (2002), S. 185–207
- [Hecking 2000] HECKING, Matthias: The SINIX Consultant - Towards a Theoretical Treatment of Plan Recognition. In: *Artificial Intelligence Review* 14 (2000), Nr. 3, S. 153–179
- [Heckner 2005] HECKNER, Markus: *Videobasierte Personenidentifikation im Fahrzeug – Design, Entwicklung und Evaluierung eines prototypischen Mensch Maschine Interfaces*, Universität Regensburg, Diplomarbeit, 2005
- [Hedicke 2002] HEDICKE, Volkmar: Multimodalität in Mensch-Maschine-Schnittstellen. In: TIMPE, Klaus-Peter (Hrsg.) ; JÜRGENSOHN, Thomas (Hrsg.) ; KOLREP-ROMETSCH, Harald (Hrsg.): *Mensch-Maschine-Systemtechnik*. Düsseldorf : Symposion Publishing, 2002, S. 203–232
- [Heisterkamp 2001] HEISTERKAMP, Paul: Linguatronic: Product-Level Speech System for Mercedes-Benz Cars. In: *Proceedings of the 1st International Conference on Human Language Technology Research*, 2001, S. 1–2
- [Helander et.al. 1997] HELANDER, Martin (Hrsg.) ; LANDAUER, Thomas (Hrsg.) ; PRABHU, Prasad (Hrsg.): *Handbook of Human-Computer Interaction*. Elsevier Science B.V., 1997
- [Hewitt et.al. 1992] HEWITT, T. ; BAECKER, R. ; CARD, S. ; CAREY, T. ; GASEN, J. ; MANTEI, M. ; PERLMAN, G. ; STRONG, G. ; VERPLANK, W.: *ACM SIGCHI Curricula for Human-Computer Interaction*. 1992
- [Hjalmarsson 2002] HJALMARSSON, Anna: *Evaluating AdApt, a mulitmodal, conversational dialogue System using PARADISE*. Stockholm, KTH Royal Institute of Technology, Diplomarbeit, November 2002
- [Hockey et.al. 2003] HOCKEY, Beth A. ; LEMON, Oliver ; CAMPANA, Ellen ; HIATT, Laura ; AIST, Gregory ; HIERONYMUS, James ; GRUENSTEIN, Alexander ; DOWDING, John: Targeted Help for Spoken Dialogue Systems: Intelligent Feedback Improves Naive User's Performance. In: *Proceedings of the 10th Conference on European Chapter of the Association for Computational Linguistics*, Association for Computational Linguistics, 2003, S. 147–154
- [Hockey und Rayner 2005] HOCKEY, Beth A. ; RAYNER, Manny: Comparison of Grammar-Based and Statistical Language Models Trained on the Same Data. In: *Proceedings of the AAAI Workshop on Spoken Language Understanding*, 2005
- [Hof 2004] HOF, Alexander: *Entwurf, Implementierung und Test eines erweiterten Hilfekonzpts für Sprachdialoge im Fahrzeug*, Universität Regensburg, Diplomarbeit, 2004

- [Hof und Hagen 2006] HOF, Alexander ; HAGEN, Eli: Help Strategies for Speech Dialogue Systems in Automotive Environments. In: ANDRE, Elisabeth (Hrsg.) ; DYBKJÆR, Laila (Hrsg.) ; MINKER, Wolfgang (Hrsg.) ; NEUMANN, Heiko (Hrsg.) ; WEBER, Michael (Hrsg.): *Perception and Interactive Technologies*. Berlin, Heidelberg, New York, Barcelona, Budapest, Hong Kong, London, Mailand, Paris, Tokyo : Springer Verlag, 2006, S. 107–116
- [Hohenner 2004] HOHENNER, Sascha: *Automatische Spracherkennung für agierende Systeme*, Universität Bielefeld, Dissertation, 2004. – URL <http://bieson.ub.uni-bielefeld.de/volltexte/2005/671/>
- [Horvitz et. al. 1998] HORVITZ, Eric ; BREESE, Jack ; HECKERMAN, David ; HOVEL, David ; KOOSROMMELSE: The Lumiere Project: Bayesian User Modeling for Inferring the Goals and Needs of Software Users. In: *Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence*, 1998, S. 256–265
- [Huber 2005] HUBER, Oswald: *Das psychologische Experiment: eine Einführung*. 4. Auflage. Bern : Huber, 2005
- [Johannson 1993] JOHANNSON, Gunnar: *Mensch-Maschine-Systeme*. Berlin, Heidelberg, New York, Barcelona, Budapest, Hong Kong, London, Mailand, Paris, Tokyo : Springer Verlag, 1993
- [Johansson 2004] JOHANSSON, Pontus: *Design and Development of Recommender Dialogue Systems*, University of Linköping, Dissertation, 2004
- [Junqua 1993] JUNQUA, Jean-Claude: The Lombard Reflex and its Role on Human Listeners and Automatic Speech Recognizers. In: *Journal of the Acoustic Society of America* 93 (1993), Nr. 1, S. 510–524
- [Kass und Finin 1988] KASS, Robert ; FININ, Tim: Modeling the User in Natural Language Systems. In: *Computational Linguistics* 14 (1988), Nr. 3, S. 5–22
- [Keller 2004] KELLER, Andreas: Dialogsysteme. In: CARSTENSEN, Kai-Uwe (Hrsg.) ; EBERT, Christian (Hrsg.) ; ENDRIS, Cornelia (Hrsg.) ; JEKAT, Susanne (Hrsg.) ; KLABUNDE, Ralf (Hrsg.) ; LANGER, Hagen (Hrsg.): *Computerlinguistik und Sprachtechnologie*. Heidelberg : Spektrum Akademischer Verlag, 2004, Kap. 5.9, S. 533–539
- [Kobsa und Wahlster 1990] KOBASA, Alfred ; WAHLSTER, Wolfgang: User Models in Dialog Systems. In: KOBASA, Alfred (Hrsg.) ; WAHLSTER, Wolfgang (Hrsg.): *User Models in Dialog Systems*. Berlin, Heidelberg, New York, Barcelona, Budapest, Hong Kong, London, Mailand, Paris, Tokyo : Springer Verlag, 1990 (Springer Series On Symbolic Computation And Artificial Intelligence), S. 4–34
- [Komatani et. al. 2003] KOMATANI, Kazunori ; ADACHI, Fumihiro ; UENO, Shinichi ; KAWAHARA, Tatsuya ; OKUNO, Hiroshi: Flexible Spoken Dialogue System based on User Models and Dynamic Generation of VoiceXML Scripts. In: *4th SIGdial Workshop on Discourse and Dialogue*, 2003, S. 87–96

- [Krahmer et.al. 1999] KRAHMER, Emiel ; LANDSBERGEN, Jan ; POUTEAU, Xavier: How to Obey the 7 Commandments for Spoken Dialogue? In: *Proceedings of the ACL/EACL Workshop on Interactive Spoken Dialog Systems*, 1999, S. 82–89
- [Krause 1988] KRAUSE, Jürgen: Adaptierbarkeit, Adaptivität, Intervenierbarkeit und Hilfesysteme / Universität Regensburg. 1988. – Forschungsbericht
- [Krause et.al. 1993] KRAUSE, Jürgen ; MITTERMAIER, Eva ; HIRSCHMANN, Astrid: The Intelligent Help System COMFOHELP - Towards a Solution of the Practicality Problem for User Modeling and Adaptive Systems. In: *User Modeling and User-Adapted Interaction* 3 (1993), Nr. 3, S. 249–282
- [Larsen 2003] LARSEN, Lars B.: Issues on the Evaluation of Spoken Dialogue Systems using Objective and Subjective Measures. In: *Proceedings of the 8th IEEE Workshop on Automatic Speech Recognition and Understanding*, 2003, S. 209–214
- [Libuda und Kraiss 2003] LIBUDA, Lars ; KRAISS, Karl-Friedrich: Dialogassistentz im Kraftfahrzeug. In: *45. Fachausschusssitzung Anthropotechnik der DGLR: Entscheidungsunterstützung für die Fahrzeug- und Prozessführung*, Oktober 2003, S. 255–270
- [Maybury und Wahlster 1998] MAYBURY, Mark ; WAHLSTER, Wolfgang: Intelligent User Interfaces: An Introduction. In: MAYBURY, Mark (Hrsg.) ; WAHLSTER, Wolfgang (Hrsg.): *Intelligent User Interfaces*. San Francisco : Morgan Kaufmann Publishers, 1998, S. 1–13
- [Mayer und Moreno 2003] MAYER, Richard ; MORENO, Roxana: Nine Ways to Reduce Cognitive Load in Multimedia Learning. In: *Educational Psychologist* 38 (2003), Nr. 1, S. 43–52
- [McTear 2002] MCTEAR, Michael: Spoken Dialogue Technology: Enabling the Conversational User Interface. In: *ACM Computing Surveys (CSUR)* 34 (2002), Nr. 1, S. 90–169
- [Mercedes Benz 2007] MERCEDES BENZ: *Die SL-Klasse in Nahaufnahme*. 2007. – URL http://www.mercedes-benz.de/content/germany/mpc/mpc_germany_website/de/home_mpc/passenger_cars/home/products/new_cars/sl-class_06/360_views.0003.html. – Letzter Zugriff am 25.06.2007
- [Miller 1956] MILLER, George: The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information. In: *The Psychological Review* 63 (1956), S. 81–97
- [Mittermaier 1995] MITTERMAIER, Eva: *Planbasierte intelligente Hilfe als Baustein eines symbiotischen Gesamtsystems*. Konstanz : UVK, 1995 (Schriften zur Informationswissenschaft)

- [Moise 2003] MOISE, Adrian: *Designing Better User Interfaces For Radiology Interpretation*, Simon Fraser University, Dissertation, 2003
- [Müller et. al. 2001] MÜLLER, C. ; GROSSMANN-HUTTER, B. ; JAMESON, A.: Rocognizing Time Pressure and Cognitive Load on the Basis of Speech: An Experimental Study. In: BAUER, Mathias (Hrsg.) ; GMYTRASIEWICZ, Piotr (Hrsg.) ; VASSILEVA, Julita (Hrsg.): *Proceedings of the 8th International Conference on User Modeling*. Berlin, Heidelberg, New York, Barcelona, Budapest, Hong Kong, London, Mailand, Paris, Tokyo : Springer Verlag, 2001, S. 24–33
- [Net et. al. 2001] NET, C. ; IYENGAR, G. ; POTAMIANOS, G. ; SENIOR, A. ; MAISON, B.: Perceptual Interfaces for Information Interaction: Joint Processing of Audio and Visual Information for Human-Computer-Interaction. In: YUAN, B. (Hrsg.) ; HUANG, T. (Hrsg.) ; TANG, X. (Hrsg.): *Proceedings of the International Conference on Spoken Language Processing* Bd. 3. Peking : Chinese Friendship Publishers, 2001, S. 11–14
- [Neuss 2001] NEUSS, Robert: *Usability Engineering als Ansatz zum Multimodalen Mensch-Maschine-Dialog*, TU München, Dissertation, 2001
- [Newell und Rosenbloom 1981] NEWELL, Allen ; ROSENBLOOM, Paul: Mechanisms of Skill Acquisition and the Law of Practice. In: ANDERSON, J. R. (Hrsg.): *Cognitive Skills and Their Acquisition*. Hillsdale, NJ : Erlbaum, 1981, S. 1–55
- [Niedermaier 2003] NIEDERMAIER, Bernhard: *Entwicklung und Bewertung eines Rapid-Prototyping Ansatzes zur multimodalen Mensch-Maschine-Interaktion im Fahrzeug*, Technische Universität München, Dissertation, 2003
- [Nielsen 1994] NIELSEN, Jakob: *Guerrilla HCI: Using Discount Usability Engineering to Penetrate the Intimidation Barrier*. 1994. – URL http://www.useit.com/papers/guerrilla_hci.htm
- [Nielsen 1997] NIELSEN, Jakob: *Usability Testing*. Kap. 46, S. 1543–1568. In: SALVENDY, Gavriel (Hrsg.): *Handbook of Human Factors and Ergonomics, Second Edition*. New York : Elsevier Press, 1997
- [Nigay und Coutaz 1993] NIGAY, Laurence ; COUTAZ, Joelle: A Design Space for Multimodal Systems: Concurrent Processing and Data Fusion. In: *Proceedings of INTERCHI '93*, ACM Press: New York, 1993, S. 172–178
- [Oviatt 1999] OVIATT, Sharon: Ten Myths of Multimodal Interaction. In: *Communications of the ACM* 42 (1999), Nr. 11, S. 74 – 81
- [Oviatt 2000a] OVIATT, Sharon: Multimodal System Processing in Mobile Environments. In: *Proceedings of the 13th annual ACM Symposium on User Interface Software and Technology*. New York : ACM Press, 2000, S. 21–30
- [Oviatt 2000b] OVIATT, Sharon: Taming Recognition Errors with a Multimodal Interface. In: *Communications of the ACM* 43 (2000), Nr. 9, S. 45–51

- [Oviatt et. al. 2001] OVIATT, Sharon ; COHEN, Philip ; WU, Lizhong ; VERGO, John ; LISBETHDUNCAN ; SUHM, Bernhard ; BERS, Josh ; HOLZMAN, Thomas ; LANDAY, Terry Winogradand J. ; LARSON, Jim ; FERRO, David: Designing the User Interface for Multimodal Speech and Gesture Applications: State-Of-The-Art Systems and Research Directions. In: CARROLL, Jennifer (Hrsg.): *Human Computer Interaction* Bd. 15. Reading, MA : Addison-Wesley Press, 2001, Kap. 19, S. 421–456
- [Oviatt et. al. 1997] OVIATT, Sharon ; DEANGELI, Antonella ; KUHN, Karen: Integration and Synchronization of Input Modes during Multimodal Human-Computer Interaction. In: *Proceedings of the Conference on Human Factors in Computing Systems*. New York : ACM Press, 1997, S. 415–422
- [Paivio 1991] PAIVIO, Allan: *Images in Mind: The Evolution of a Theory*. New York London Toronto Sydney Tokyo Singapore : Harvester Wheatsheaf, 1991
- [Peckham 1991] PECKHAM, Jeremy: Speech Understanding and Dialogue over the Telephone: an Overview of the ESPRIT SUNDIAL Project. In: *Proceedings of the Speech and Natural Language Workshop*, Pacific Grove, CA, 1991, S. 14–27
- [Piechulla et. al. 2002] PIECHULLA, Walter ; MAYSER, Christoph ; GEHRKE, Helmar ; KÖNIG, Winfried: Echtzeit-Fahrerbeanspruchungsschätzung. In: HAMMWÖHNER, Rainer (Hrsg.) ; WOLFF, Christian (Hrsg.) ; WOMSER-HACKER, Christa (Hrsg.): *Information und Mobilität. Proceedings des 8. Internationalen Symposiums für Informationswissenschaft*. Konstanz : UVK, 2002, S. 365–368
- [Ramachandran und Young 2005] RAMACHANDRAN, Ashwin ; YOUNG, R. M.: Providing Intelligent Help Across Applications in Dynamic User and Environment Contexts. In: *IUI '05: Proceedings of the 10th International Conference on Intelligent User Interfaces*. New York, NY, USA : ACM Press, 2005, S. 269–271
- [Rasch et. al. 2004] RASCH, Björn ; FRIESE, Malte ; HOFMANN, Wilhelm ; NAUMANN, Ewald: *Quantitative Methoden*. Berlin, Heidelberg, New York, Barcelona, Budapest, Hong Kong, London, Mailand, Paris, Tokyo : Springer Verlag, 2004
- [Reuther 2003] REUTHER, Achim: *useML - Systemetische Entwicklung von Maschinenbediensystemen mit XML*, Universität Kaiserslautern, Dissertation, 2003
- [Ritter und Schooler 2002] RITTER, Frank E. ; SCHOOLER, Lael J.: The Learning Curve. In: *International Encyclopedia of the Social and Behavioral Sciences*. Amsterdam : Pergamon, 2002, S. 8602–8605
- [Rubin und Wenzel 1996] RUBIN, David ; WENZEL, Amy: One Hundred Years of Forgetting: A Quantitative Description of Retention. In: *Psychological Review* 103 (1996), Nr. 4, S. 734–760

- [Rubin 1994] RUBIN, Jeffrey: *Handbook of Usability: How to Plan, Design and Conduct Effective Tests*. Toronto : John Wiley & Sons, Inc., 1994
- [Sagstetter 2007] SAGSTETTER, Stefan: *Evaluation eines Hilfesystems für Sprachdialoge im Fahrzeug*, Universität Regensburg, Diplomarbeit, 2007
- [Salmen 2003] SALMEN, Angelika: *Multimodale Menüausgabe im Fahrzeug*, Universität Regensburg, Dissertation, 2003
- [Scheufler 2002] SCHEUFLER, Ingrid: Qualitative Bestimmung der situativen Fahrerbelastung. In: *38. BDP-Kongress für Verkehrspsychologie*, 2002
- [Schukat-Talamazzini 1995] SCHUKAT-TALAMAZZINI, Ernst G.: *Automatische Spracherkennung*. Braunschweig : Vieweg Verlag, 1995
- [Schütz und Schäfer 2002] SCHÜTZ, W. ; SCHÄFER, R.: Towards More Realistic Modeling of a User's Evaluation Process. In: *Personalization of the Mobile World*, 2002, S. 91–98
- [Schweigert 2002] SCHWEIGERT, Manfred: *Fahrerblickverhalten und Nebenaufgaben*, Technische Universität München, Dissertation, 2002
- [Shneiderman 1997] SHNEIDERMAN, Ben: *Designing the User Interface: Strategies for Effective Human-Computer Interaction*. 3. Ausgabe. Reading, Massachusetts : Addison-Wesley, 1997
- [Stangl 1997] STANGL, Werner: *Das Vergessen*. <http://paedpsych.jk.uni-linz.ac.at/INTERNET/arbeitsblaetterord/LERNTECHNIKORD/Vergessen.html>. 1997. – Letzter Zugriff: 21.01.2007
- [Stephanidis et.al. 1999] STEPHANIDIS, Constantine ; AKOUMIANAKIS, Demosthenes ; PARAMYTHIS, Alexandros: *Coping with diversity in HCI: Techniques for adaptable and adaptive interaction*, Tutorial 11. 1999. – URL <http://www.ics.forth.gr/proj/at-hci/publications/tutorials.html>. – 8th International Conference on Human-Computer Interaction
- [Sun Developer Network 1998] SUN DEVELOPER NETWORK: *JSRG Grammar Format Specification*. 1998. – URL <http://java.sun.com/products/java-media/speech/forDevelopers/JSRG/>. – Letzter Zugriff am 10.01.2007
- [Totzke 2001] TOTZKE, Ingo: Die Humanisierung multimodaler HMI – die Bedeutung der Modalität und Situation der Mensch-Maschine-Interaktion. In: *E-Car Infotainmentplattform & Telematikdienste für das Multimedia Auto*, 2001
- [Totzke et.al. 2003a] TOTZKE, Ingo ; KRÜGER, Hans-Peter ; HOFMANN, Michael ; MEILINGER, Tobias ; RAUCH, Nadja ; SCHMIDT, Gerrit: Kompetenzerwerb für Informationssysteme / IZVW Würzburg. September 2003. – Forschungsbericht

- [Totzke et. al. 2003b] TOTZKE, Ingo ; MEILINGER, Tobias ; KRÜGER, Hans-Peter: Erlernbarkeit von Menüsystemen im Fahrzeug - mehr als nur eine Lernkurve. In: *VDI-Berichte: Der Fahrer im 21. Jahrhundert*. Düsseldorf : VDI-Verlag, 2003
- [Vollrath und Totzke 2000] VOLLRATH, Mark ; TOTZKE, Ingo: *In-Vehicle Communication and Driving: An Attempt to Overcome their Interference*. 2000
- [W3C 2004] W3C: *Voice Extensible Markup Language (VoiceXML)*. 2004. – URL <http://www.w3.org/TR/voicexml20/>
- [W3C 2005] W3C: *EMMA: Extensible MultiModal Annotation Markup Language*. 2005. – URL <http://www.w3.org/TR/emma/>
- [Wahlster 1999] WAHLSTER, Wolfgang: Sprachtechnologie im Alltag. In: *Alltag der Zukunft - Informationstechnik verändert unser Leben*, 1999
- [Walker et. al. 1997] WALKER, Marilyn ; LITMAN, Diane ; KAMM, Candace ; ABELLA, Alicia: PARADISE: A Framework for Evaluating Spoken Dialogue Agents. In: *Proceedings of the eighth conference on European chapter of the Association for Computational Linguistics*. Morristown, New Jersey : Association for Computational Linguistics, 1997, S. 271–280
- [White 2003] WHITE, Geoffrey: Forgetting Functions. In: *Animal Learning & Behavior* 29 (2003), Nr. 3, S. 193–207
- [Wikipedia 2007] WIKIPEDIA: *A*-Algorithmus*. 2007. – URL http://de.wikipedia.org/wiki/A*-Algorithmus. – Letzter Zugriff am 06.07.2007
- [Wirth 2002] WIRTH, Thomas: *Die magische Zahl 7 und die Gedächtnisspanne*. 2002. – URL <http://www.kommdesign.de/texte/gedaechtnisspanne.htm>
- [Zimbardo und Gerrig 1999] ZIMBARDO, Philip ; GERRIG, Richard ; HOPPE-GRAF, Siegfried (Hrsg.) ; ENGEL, Irma (Hrsg.): *Psychologie*. Berlin, Heidelberg, New York, Barcelona, Budapest, Hong Kong, London, Mailand, Paris, Tokyo : Springer Verlag, 1999
- [Zue und Glass 2000] ZUE, Victor ; GLASS, James: Conversational Interfaces: Advances and Challenges. In: *Proceedings of the IEEE* Bd. 88, 2000, S. 1166–1180
- [Zue et. al. 2000] ZUE, Victor ; SENEFF, Stephanie ; GLASS, James ; POLIFRONI, Joseph ; PAO, Christine ; HAZEN, Timothy J. ; HETHERINGTON, Lee: Jupiter: A Telephone-Based Conversational Interface for Weather Information. In: *IEEE Transactions on Speech and Audio Processing* 8 (2000), Nr. 1, S. 100–112
- [Zühlke 2002] ZÜHLKE, Detlev: Useware - Herausforderung der Zukunft. In: *atp - Automatisierungstechnische Praxis* 44 (2002), Nr. 9, S. 73–77

Anhang A

Abkürzungsverzeichnis

Abkürzungsverzeichnis

ASE	Automatische Spracherkennung
ASR	Automatic Speech Recognition
AVM	Attribute-Value-Matrix
BSD	Between-Subject-Design
CID	Central Information Display
DCT	Domain Concept Tree
EH	Evidential Horizon
EM	Execution Manager
EMMA	Extensible Multimodal Annotation Markup Language
FAS	Fahrerassistenzsystem
FIS	Fahrerinformationssystem
GBM	Grammar Based Language Model
GDML	Generic Dialogue Modeling Language
GUI	Graphical User Interface
HMM	Hidden-Markov-Modell
HUD	Head-Up-Display
JSGF	Java Speech Grammar Format
KZG	Kurzzeitgedächtnis
LEL	Lumière Events Language
LLP	Logical Language For Planning
MATCH	Multimodal Access to City Help

MAUT	Multi-Attribute-Utility-Theorie
MMI	Mensch-Maschine-Interaktion
MMS	Mensch-Maschine-System
NASA-TLX	NASA Taskload Index
OOV	Out-of-Vocabulary
PARADISE	Paradigm for Dialogue System Evaluation
PTT	Push-to-Talk
SALT	Speech Application Language Tags
SDS	Sprachdialogsystem
SLM	Statistic Language Model
SMEQ	Subjective Mental Effort Questionnaire
SWAT	Subjective Workload Assessment Technique
TH	Targeted Help
TTS	Text-to-Speech
VP	Versuchsperson
WSD	Within-Subject-Design

Anhang B

Intentionen Pool

[illegible]

Anhang C

Testdrehbuch

Begrüßung	
alle	Abholen VP am Eingang Begrüßung
Einverständniserklärung / Geheimhaltung	
alle	VP liest und versteht Erklärung Unterschrift durch VP
Betreten Usability-Labor	
VP, Beifahrer	
Start Fahrsimulation	
Technik	Starten der Fahrsimulation
Start Video-Aufzeichnung	
Technik	Starten der Videoaufzeichnung
Instruktion	
Beifahrer	<p>Ziel dieses Usability-Testes ist die Prüfung eines frühen Prototypen für ein neues Spracheingabekonzept im Fahrzeug. Wir möchten feststellen, inwieweit die bisher implementierten Funktionen den Bedürfnissen von Ihnen als Nutzer entsprechen.</p> <p>Sie sitzen nun in einer Sitzkiste. Wie Sie sehen ist der Innenraum dem eines normalen Fahrzeuges nachempfunden.</p> <p>Vorne sehen Sie drei große Bildschirme. Auf diesen wird eine Fahrsimulation angezeigt.</p> <p>Im Innenraum sind drei Kameras mit Mikrofon angebracht. Damit wollen wir Ihre Benutzung des Systems aufzeichnen, um später leichter Auswertungen und Aussagen zum System machen zu können.</p> <p>Mit dem erwähnten Fahrsimulator werden Sie wie im normalen Fahrzeug fahren. Das ist nötig, weil wir die Benutzung der Sprachschnittstelle im normalen Einsatz auf der Straße simulieren möchten.</p> <p>Weiterhin werde ich Ihnen einige typische Aufgaben stellen, die im Alltag bei der Bedienung der Sprachschnittstelle auftreten können. Ich werde Ihnen diese Aufgaben jeweils vorlesen, damit Sie sich auf das Fahren konzentrieren können.</p>
VP	ALLES OK?
Einweisung Fahrsimulation	
Beifahrer	Nun wollen wir uns dem Fahrsimulator widmen. Sie haben nun die Möglichkeit sich mit dem Fahrsimulator vertraut zu

	<p>machen. Bei der Fahrsimulation handelt es sich um eine Folgefahrt. D.h. Sie müssen dem vorausfahrenden Fahrzeug folgen. Ändert bspw. das vorausfahrende Fahrzeug die Geschwindigkeit, so müssen Sie Ihre Geschwindigkeit ebenfalls anpassen.</p> <p>...</p>
VP	3 Minuten Einfahren und Gewöhnung an Simulator
VP	3 Minuten Referenzfahrt
Beifahrer	<p>Sehr gut. Ich würde sagen, dass wir nun mit den Aufgaben beginnen. In der ersten Aufgabe werden Sie den Prototypen kennen lernen und einige grundlegende Handlungen durchführen, damit Sie schon mal sehen, wie die Bedienung funktioniert.</p> <p>Eine Sache noch vorweg. Ich hoffe, Sie fühlen sich nicht wie bei einem Test. Denn, Ziel dieser Sitzung ist es, den Prototypen zu testen und bei diesem Schwachstellen zu finden; wir wollen nicht Sie testen.</p> <p>Sollte Ihnen etwas auffallen, sollten Situationen unangenehm sein oder Sie ein Umstand zufrieden stellen bzw. unzufrieden machen, dann bitte ich Sie, mir dies einfach mitzuteilen.</p>
Aufgabe 1 – Hauptmenü	
Beifahrer	<p>Mit Aufgabe 1 möchte ich Ihnen kurz den Prototypen samt Sprachschnittstelle zeigen, sodass Sie einen ersten Einblick gewinnen können.</p> <p>Wenn Sie das Display in der Mitte betrachten, dann sehen Sie zunächst das Hauptmenü. Dieses Menü wird in den folgenden Aufgaben immer als Ausgangspunkt dienen.</p> <p>Sie können das System mit dem Controller auf der Mittelkonsole bedienen. Den Controller können Sie drehen, nach oben unten links und rechts schieben und darauf drücken.</p> <p>Es gibt eine so genannte Push-To-Talk-Taste. Diese Taste befindet sich am rechten Teil des Multifunktionslenkrades. Wenn Sie diese Taste drücken, können Sie nachfolgend ein Sprachkommando eingeben. Durch nochmaliges Drücken können Sie die Spracheingabe wieder beenden.</p> <p>Sie möchten nun zum ersten Mal die Sprachschnittstelle benutzen. Über das Kommando <i>Hilfe</i> erhalten Sie erste Infor-</p>

	mationen zum System. Betätigen Sie nun die Pust-To-Talk-Taste. Sprechen Sie <i>Hilfe</i> .
VP	„Hilfe“
Beifahrer	PTT nochmals drücken lassen nach Ausgabe
Beifahrer	Mit dem Kommando <i>Sprachoptionen</i> finden Sie heraus, welche Möglichkeiten Sie im aktuellen Menü haben. Sprechen Sie jetzt <i>Sprachoptionen</i> .
VP	„Sprachoptionen“
Beifahrer	PTT nochmals drücken lassen nach Ausgabe
Beifahrer	<p><i>Sprachoptionen</i> können Sie immer dann verwenden, wenn Sie die im aktuellen Untermenü gerade möglichen Kommandos nicht kennen und wissen möchten, welche Möglichkeiten bzw. Optionen sie haben. Bitte machen Sie das auch. D.h. wenn Sie zum ersten Mal in ein neues Menü gelangen und keinen Überblick über die möglichen Befehle haben, dann verwenden Sie einfach das Kommando <i>Sprachoptionen</i>.</p> <p>-</p> <p>Immer dann, wenn Sie das Kommando <i>Sprachoptionen</i> verwenden, haben Sie nicht nur die Möglichkeit einen Menüpunkt durch das Sprechen auszuwählen, sondern auch durch die Betätigung des Controllers. Sofort beim Hören Ihrer gewünschten Auswahl, können Sie den Controller drücken, wodurch Sie sofort eine Ebene weiter kommen.</p>
Beifahrer	<p>Das Kommando <i>Zurück</i> funktioniert wie der Back/Zurück-Button im Browser.</p> <p>Sie können jederzeit, egal wo Sie sich befinden, das Hauptmenü aufrufen, indem Sie das Kommando <i>Hauptmenü</i> benutzen. Dadurch gelangen Sie wieder zur ursprünglichen Ausgangsposition zurück.</p> <p>Es kann auch vorkommen, dass Sie der Spracherkenner nicht eindeutig versteht. Sollte das der Fall sein, dann werden Ihnen zwei bis drei alternative Möglichkeiten zur Auswahl gestellt, die sie dann mit <i>Eintrag 1</i>, <i>Eintrag 2</i> oder <i>Eintrag 3</i> verwenden können.</p>
Beifahrer	ALLES OK? NOCH FRAGEN?
Aufgabe 2 – Jemanden anrufen	
Beifahrer	Sie möchten Markus Abel aus Ihrem Telefonbuch anrufen. Im Telefonbuch angekommen können Sie durch das Kommando <i>Sprachoptionen</i> herausfinden, welche Möglichkeiten Sie haben, diesen Anruf zu erledigen.
Beifahrer	Hinweis auf Kommando „ <i>Auswahl</i> “ bzw. Controller nach Links

VP	<p>Telefon – Telefonbuch – Sprachoptionen</p> <p>a) Eintrag Markus Abel anrufen</p> <p>b) Nachname Vorname anrufen / Vorname Nachname anrufen</p> <p>c) Eintrag x anrufen</p>
Aufgabe 3 – Wahlwiederholung	
Beifahrer	<p>Ich wähle nun eine Telefonnummer, die sich nicht im Adressbuch befindet.</p> <p>Wahl von irgendwas und sofort wieder auflegen.</p> <p>Startpunkt: Menü Telefon</p>
Beifahrer	<p>Nachdem der Anruf beendet ist, wollen Sie diese Nummer erneut anrufen.</p> <p>Sie möchten die Nummer während der Fahrt nicht nochmals eintippen. Finden Sie deshalb eine andere Möglichkeit die eben gewählte Nummer nochmals anzurufen.</p>
VP	Telefon – Sprachoptionen – Weiter – Wahlwiederholung – Eintrag 1 anrufen
Beifahrer	Falls Controller nach <i>Sprachoptionen</i> - Frage: Wie hat Ihnen die Möglichkeit der Auswahl durch den Controller zugesagt?
Beifahrer	<p>Frage: Fällt Ihnen nach dem Kommando Sprachoptionen während der akustischen Ausgabe auf dem Bildschirm etwas aufgefallen?</p> <p>-</p> <p>Wenn ja, dann: Wie nützlich fanden Sie das? Hilft Ihnen die gleichzeitige visuelle Darstellung?</p> <p>-</p> <p>Wenn nein, dann: Verwenden Sie nochmals das Kommando Sprachoptionen und schauen Sie, sofern es der Verkehr gestattet auch kurz auf den Bildschirm.</p> <p>Ist Ihnen etwas aufgefallen?</p> <p>-</p> <p>Frage: Fanden Sie die gleichzeitige visuelle Ausgabe nützlich?</p> <p>-</p> <p>Frage: Hat Ihnen die Ausgabe auf zwei Wegen geholfen, dass System besser zu bedienen?</p>
Aufgabe 4 – Nachricht löschen	
Beifahrer	Im Nachrichtenspeicher befindet sich eine SMS mit einer unbekannten Nummer; also kein Name aus dem Adressbuch sondern nur eine Ziffernfolge. Sie möchten nun diese SMS löschen.
VP	Telefon – Nachrichten – Sprachoptionen – Nachricht löschen

	schen
Beifahrer	Löschen Sie eine weitere SMS bitte nun nur mit Hilfe des Controllers.
VP	Löschen mit Controller
Beifahrer	Frage: Welche Art der Bedienung hat Ihnen in diesem Fall besser gefallen? - Frage: Können Sie sich Situation vorstellen, wann Sie eher die erste, wann eher die zweite Art der Bedienung wählen würden?
Aufgabe 5 – Kartenansicht wechseln	
Beifahrer	Sie möchten die verschiedenen Möglichkeiten der Kartenansicht kennen lernen. Dazu müssen Sie zunächst herausfinden, wo Sie die Darstellung der Karte ändern können. Wenn Sie das Menü gefunden haben, möchten Sie die Karte in Fahrtrichtung ausgerichtet sehen.
VP	Navigation – Karte – Sprachoptionen – Ansichten –Karte fahrtweisend
Beifahrer	Nun soll die Karte immer in Richtung Norden zeigen.
VP	Navigation – Karte – Sprachoptionen – Ansichten –Karte nordweisend
Beifahrer	Ist Ihnen noch eine weitere Möglichkeit der Kartenansicht aufgefallen / im Gedächtnis?
VP	Antwort
Beifahrer	Wie schwierig finden Sie es, die Kartenansicht zu ändern und gleichzeitig dem Verkehr zu folgen?
Aufgabe 6 – Maßstab wählen	
Beifahrer	Startpunkt: Menü Navigation
Beifahrer	Auch der Maßstab der Karte ist einstellbar. Bitte finden Sie heraus, wie das funktioniert.
VP	Navigation – Karte – Sprachoptionen – Maßstab wählen
Beifahrer	Testen Sie nun der Reihe nach die Maßstäbe 500m, 2km, 5km
Pause	
Technik	Neustart Fahrsimulation
Alle	Getränke, Kekse, Toilette
Einfahren	
VP	2 Minuten Einfahren und zusätzliche Referenz
Aufgabe 7 – Routenkriterien	
Beifahrer	Sie wollen die Streckenführung nun beeinflussen, als bspw. ein Kriterium wie die Benutzung von Autobahnen. Finden Sie das dazu nötige Menü.
VP	Navigation – Karte – Routenkriterien – Sprachoptionen Mautstrecken vermeiden aus

Beifahrer	Mautstrecken dürfen verwendet werden
VP	Maustrecken vermeiden aus
Beifahrer	Frage: Wie würden Sie die Benutzung von Fähren einstellen?
VP	Antwort
Aufgabe 8 - Senderwahl	
Beifahrer	<p>Bisher haben Sie durch die Verwendung des Kommandos Sprachoptionen immer die so genannte kontextabhängige Hilfe verwendet. D.h., dass Sie für das aktuelle Untermenü, also abhängig von jeweiligen Kontext, die möglichen Sprachkommandos erhalten haben.</p> <p>Es ist im Prototypen auch eine kontextunabhängige Hilfe vorgesehen. Sie können jederzeit, egal in welchem Menü Sie sich befinden, fragen, wie Sie eine bestimmte Funktion erreichen und ausführen können.</p> <p>Das sieht zum Beispiel so aus: Sie sprechen das Kommando <i>Einführung</i>. Nach der Sprachausgaben sprechen Sie „<i>Wie kann ich telefonieren?</i>“. Auf diese Frage folgt eine Sprachausgabe, die Ihnen die passende Information zur Verfügung stellt.</p>
Beifahrer	Sie möchten nun den Sender Bayern 3 hören. Bitte finden Sie nun mit einer Frage den Weg zum Ziel.
VP	<p>Radio – Sender wählen – [Sprachoptionen] – Bayern 3</p> <p>Auswahl über Frage: “Wie kann ich Radio hören“ o. ä.</p>
Aufgabe 9 – Verkehrsfunk an/aus	
Beifahrer	Starpunkt: Menü Radio
Beifahrer	Sie möchten nun den Verkehrsfunk anstellen.
VP	<p>Radio</p> <p>[Sprachoptionen]</p> <p>Verkehrfunk an</p>
Beifahrer	Sie möchten den Verkehrsfunk nun wieder deaktivieren.
VP	<p>Radio</p> <p>[Sprachoptionen]</p> <p>Verkehrsfunk aus</p>
Aufgabe 10 – CD-Titel auswählen	
Beifahrer	Sie möchten nun eine CD hören. Wählen Sie Track 5 von CD 3 im CD-Wechsler.
VP	<p>CD Multimedia –</p> <p>Sprachoptionen –</p> <p>CD 3 Titel 5 CD3 Eintrag 5</p>
Beifahrer	Nun möchten Sie zu Eintrag 8 auf derselben CD springen.

VP	CD Multimedia – Sprachoptionen – Eintrag 8 Titel 8
Aufgabe 11 – Vorheriger/Nächster Titel	
Beifahrer	Sie möchten die Titel der Reihe nach hören. Spiele Sie einen Titel nach dem anderen ab.
VP	CD Multimedia – CD Laufwerk – Sprachoptionen – Vorheriger Titel – Nächster Titel
Aufgabe 12 – Klangeinstellung	
Beifahrer	Sie möchte nun wiederum mit eine Frage wissen, wo Sie die Bässe verstellen können. Bitte fragen Sie nun nach der Möglichkeit der Bässeverstellung.
VP	Antwort
Beifahrer	Wie würden Sie fragen, wenn Sie eine CD anhören möchten?
VP	Antwort
Direkte Nachfragen	
Beifahrer	Falls Fragen aufgetaucht sind, diese sofort im Simulator noch stellen
Verlassen Usability Labor	
VP, Beifahrer	Verlassen des Simulators
Alle	Getränke, Kekse, Toilette
Fragebogen	
VP	Ausfüllen; Fragen beantworten
Verabschiedung	
Alle	

Anhang D

Fragebogen

Alle Ihre Angaben werden vertraulich behandelt und nicht an Dritte weitergegeben.
Die Auswertung Ihrer Antworten erfolgt anonym. Bitte geben Sie ehrliche und objektive Antworten.

Allgemeine Angaben:

Geschlecht: Männlich ☐ Weiblich ☐

Alter: _____

Höchster Abschluss:

Keiner	Hauptschule	Realschule	Berufsausbildung
Meister	Univ./FH-Abschluss	Promotion	Anderer: _____

Beruf: _____

Jahre mit Führerschein: _____ Jahre

Fahrleistung pro Jahr: 0 bis 10.000 km ☐ 10.001 bis 25.000 km ☐ mehr als 25.000 km ☐

Fahrleistung Gesamt: 0 bis 100.000 km ☐ mehr als 100.000 km ☐

Sie sind: Linkshänder ☐ Rechtshänder ☐ Beidhändig ☐

BMW-Fahrer: ja ☐ nein ☐

Umgang mit technischen Geräten

Wie häufig benutzen Sie einen PC?

- | | |
|---|---|
| <input type="checkbox"/> Täglich | <input type="checkbox"/> 1-3x Monatlich |
| <input type="checkbox"/> Mehrmals pro Woche | <input type="checkbox"/> Weniger als 1x pro Monat |
| <input type="checkbox"/> 1x Wöchentlich | <input type="checkbox"/> Nie |

Hinweis:

Im folgenden werden Sie immer wieder auf sog. Likert-Skalen treffen. Ein Beispiel:

Wie wichtig ist es für Sie einen hohen Lebensstandard zu haben?

Sehr wichtig ____ _X_ ____ unwichtig

Sie haben dann die Möglichkeit Ihre Einstellung oder Meinung zu einer Frage abgestuft (6 Stufen) durch Ankreuzen anzugeben. Im obigen Beispiel hätte ein Kreuz ganz links die Bedeutung, dass Ihnen ein hoher Lebensstandard sehr wichtig ist, ein Kreuz äußerst rechts, dass Ihnen ein hoher Lebensstandard nichts bedeutet.

Wie schätzen Sie Ihre Erfahrung im Umgang mit Computern ein?

Sehr gut ____ Sehr schlecht

Schätzen Sie sich als technisch begabt ein?

Ja, sehr ____ Nein, gar nicht

Empfinden Sie den Einfluss des Computers als positiv oder negativ für Ihre tägliche Arbeit?

Positiv ____ Negativ

Verwenden Sie im Allgemeinen das beigelegte Handbuch, wenn Sie ein neues technisches Gerät einsetzen?

Immer ____ nie

Welche Sprachanwendungen haben Sie schon verwendet und wie oft verwenden Sie diese?

Diktiersystem	Sehr oft ____	Nie
Telefonbuchsystem	Sehr oft ____	Nie
Auskunftssystem (Kino, Flughafen, Bahnhof)	Sehr oft ____	Nie
Mobiltelefon	Sehr oft ____	Nie
Auto	Sehr oft ____	Nie
Sonstige: _____	Sehr oft ____	Nie

Kennen Sie iDrive, das Anzeige Bedienkonzept von BMW? ja ☐ nein ☐

Wenn nein, dann weiter mit der nächsten Seite.
Wenn ja, bitte noch folgende vier Fragen beantworten.

Wie häufig benutzen Sie iDrive?

Sehr oft ____ Nie

Wie schätzen Sie Ihre Erfahrung / Ihre Kenntnisse im Umgang mit iDrive ein?

Sehr gut ____ Sehr schlecht

Wie häufig benutzen Sie die Spracheingabekomponente von iDrive?

Sehr oft ____ Nie

Wie schätzen Sie Ihre Erfahrung im Umgang mit der Sprachbedienung im Fahrzeug ein?

Sehr gut ____ Sehr schlecht

Fragen zur Belastung durch die Benutzung des Prototypen

Kreuzen Sie aufgrund Ihrer Erfahrung beim Lösen der Aufgaben jeweils den Punkt auf der Skala an, durch den Ihre Belastung am Ehesten wiedergegeben wird. Durch das untenstehende Beispiel für den Begriff Langeweile wird ziemlich große Langeweile während des Tests angezeigt.

z.B. Langeweile																
Gering															X	Hoch

Geistige Anforderung: In welchem Maß stellte die Aufgabe eine geistige Anforderung, also Denken, Entscheiden, Beobachten, dar?

Gering																Hoch
--------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	------

Körperliche Anforderung: Wie viel körperliche Aktivität, also Lenken, Bremsen etc. ist erforderlich?

Gering																Hoch
--------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	------

Zeitliche Anforderung: Welchen Zeitdruck empfinden Sie aufgrund der Geschwindigkeitsanforderungen, die die Aufgabe stellt?

Hoch																Gering
------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--------

Aufgabenerfüllung: Wie zufrieden sind Sie mit dem Grad der Aufgabenerfüllung, den Sie erreicht haben?

Hoch																Gering
------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--------

Anstrengung: Wie sehr mussten Sie sich insgesamt anstrengen (geistig, körperlich), um diese Aufgabe zu bewältigen?

Gering																Hoch
--------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	------

Unzufriedenheit: In welchem Maß fühlten Sie sich bei der Erfüllung dieser Aufgabe unzufrieden?

Hoch																Gering
------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--------

Bitte entscheiden Sie nun pro Wortpaar, welcher Faktor größeren Einfluss auf Ihre Belastung / Ihr Befinden während der Durchführung der Aufgaben hatte und kreuzen Sie diesen an. Im Beispiel ist Langeweile angekreuzt, mit der Bedeutung, dass Langeweile für Sie wichtiger war als Aufregung.

Beispiel:

	Langeweile	Aufregung	
	Zeitliche Anforderung	Anstrengung	
	Unzufriedenheit	Anstrengung	
	Aufgabenerfüllung	Unzufriedenheit	
	Geistige Anforderung	Anstrengung	
	Unzufriedenheit	Geistige Anforderung	
	Aufgabenerfüllung	Geistige Anforderung	
	Körperliche Anforderung	Zeitliche Anforderung	
	Zeitliche Anforderung	Geistige Anforderung	
	Körperliche Anforderung	Unzufriedenheit	
	Geistige Anforderung	Körperliche Anforderung	
	Körperliche Anforderung	Aufgabenerfüllung	
	Anstrengung	Körperliche Anforderung	
	Zeitliche Anforderung	Unzufriedenheit	
	Aufgabenerfüllung	Zeitliche Anforderung	
	Anstrengung	Aufgabenerfüllung	

Die Aufgaben waren realistisch.

Völlig falsch ____ Völlig richtig ____

Wie empfanden Sie die gestellten Aufgaben als solche (Komplexität, Verständlichkeit, ...)?

Sehr schwierig ____ Sehr leicht ____

War das Fahren als solches anstrengend (Verkehrsaufkommen, unbekannte Strecke, ...)?

Sehr anstrengend ____ Gar nicht anstrengend ____

Wie leicht/schwierig war die Bewältigung der Aufgaben während der Fahrt?

Sehr schwierig ____ Sehr leicht ____

War das Fahren hinderlich für die Bewältigung der Aufgaben?

Ja, sehr ____ Nein, gar nicht ____

Die Bedienung des Systems / die Lösung der Aufgaben hat mein Fahrverhalten negativ beeinflusst.

Völlig richtig ____ Völlig falsch ____

Würden Sie die gestellten Aufgaben normalerweise während der Fahrt erledigen oder dazu kurz anhalten?

Im Stand ____ Während der Fahrt ____

Nennen Sie, soweit möglich, Beispiele von den gestellten Aufgaben, die Sie lieber im Stand erledigen würden.

Hilfefunktionalität

Immer dann wenn Sie das Kommando „Sprachoptionen“ verwendet haben, dann haben Sie das Hilfesystem in unserem Prototypen genutzt. Folgende Fragen beziehen sich genau auf diese Dialogschritte.

Die Hilfeausgaben passen sich automatisch Ihrem Verhalten und Ihren Kenntnissen (Länge der Dialoge, Verwendung von Beispielen) an. Befürworten Sie generell eine derartige Vorgehensweise?

Ja, sehr ____ Nein, gar nicht ____

Bitte begründen Sie Ihre Auswahl:

Wie beurteilen Sie die Länge der akustischen Hilfeausgaben?

Zu lang ____ Zu kurz

Ist Ihnen im Laufe des Tests eine Veränderung hinsichtlich der Länge oder der Inhalte der Hilfe aufgefallen?

ja ☐ nein ☐

Wenn ja, auf was führen Sie diese Veränderung zurück?

Sind Ihnen während der akustischen Hilfeausgaben auf dem Display Veränderungen aufgefallen?

ja ☐ nein ☐

Wenn ja, welche?

Empfanden Sie die zusätzliche graphische Anzeige der Hilfeausgaben als nützlich?

Ja, sehr ____ Nein, gar nicht

Die Auswahl der Sprachoptionen kann sowohl durch Sprache als auch durch den Controller erfolgen.

War diese Möglichkeit für Sie persönlich von Nutzen?

Ja, sehr ____ Nein, gar nicht

Wie nützlich war das Hilfesystem für die Erfüllung der Aufgaben?

Sehr nützlich ____ Gar nicht nützlich

Das Kommando „Sprachoptionen“ wird immer dann benutzt, wenn Befehle in einem Kontext nicht bekannt sind. Traten bei Ihnen nach der Benutzung des Kommandos „Sprachoptionen“ noch Unklarheiten auf?

ja ☐ nein ☐

Wenn ja, erinnern Sie sich noch an die betreffenden Aufgaben?

Waren die Sprachausgaben in Ihrer Länge angemessen?

Zu lang ____ Zu kurz

Waren die Hilfeanweisungen klar verständlich?

Verständlich ____ Nicht verständlich ____

Wie empfanden Sie die Bedienung des Systems?

Schwer zu bedienen ____ Leicht zu bedienen ____

Konnten die Hilfeaushaben Ihnen bei der Bedienung des Systems helfen?

Ja, immer ____ Nein, gar nicht ____

Denken Sie, dass die akustischen Hilfeaushaben Sie beim Lernen der Kommandos entscheidend unterstützt haben?

Ja, sehr ____ Nein, gar nicht ____

Welchen Einfluss hatten die Sprachoptionen auf Ihren Lernerfolg?

Sehr großen Einfluss ____ Sehr kleinen Einfluss ____

Wenn Sie die Wahl haben: Würden Sie das Lernen der möglichen Spracheingabe-Kommandos durch das Lesen eines Handbuches oder durch die Hilfefunktion des Systems (Kommando „Sprachoptionen“) bevorzugen?

Handbuch ☐ Hilfefunktion ☐

Kontextunabhängige Hilfe

Sie haben die kontextunabhängige Hilfe benutzt, wenn Sie eine Frage der Art „Wie kann ich telefonieren?“ gestellt haben.

Wie nützlich ist für Sie diese Art der Hilfe?

Sehr nützlich ____ Gar nicht nützlich ____

Waren die Antworten der kontextunabhängigen Hilfe sinnvoll und hilfreich?

Ja, sehr ____ Nein, gar nicht ____

Wie empfanden Sie die Länge der Antworten der kontextunabhängigen Hilfe?

Sehr lang ____ Sehr kurz ____

Weitere Fragen zum Spracheingabesystem

Welche Funktionen waren für Sie störend bzw. haben Sie bei der Bedienung verunsichert?

Wie schätzen Sie Ihren gewonnenen Überblick über die Möglichkeiten des Prototypen ein? War es Ihnen durch die gebotene Hilfe möglich einen guten Überblick über die Möglichkeiten des Prototypen zu gewinnen?

Ja, guter Überblick ____ Nein, schlecht Überblick

Haben Sie immer alle Informationen und Funktionen dort gefunden, wo Sie diese erwartet haben?

Ja, immer ____ Nein, nie

In der Zukunft könnte es möglich sein, auch TV-Sender zu empfangen und diese über das Display anzuzeigen. Stellen Sie sich folgende Liste vor:

1: ARD

2: RTL

3: SAT 1

4: PRO 7

5: Neuer Sender; Name unbekannt

6: Neuer Sender; Name unbekannt

Mit welchem Kommando / welchen Kommandos würden Sie den unterstrichenen Sender aufrufen?

Mit welchem Kommando / welchen Kommandos würden Sie nach Aufruf dieses Senders auf den Sender wechseln, der eine Zeile höher in der Liste steht?

Benutzerzufriedenheit (PARADISE Framework)

Das Sprachsystem war sehr benutzerfreundlich

Stimme zu ____ Stimme nicht zu

Ich konnte alle Aufgaben ohne Probleme bewältigen

Stimme zu ____ Stimme nicht zu

Das Sprachsystem war nicht benutzerfreundlich

Stimme zu ____ Stimme nicht zu

Ich habe mich während der Bedienung sehr geärgert

Stimme zu ____ Stimme nicht zu

Die Sprachsteuerung im Fahrzeug ist sehr nützlich

Stimme zu ____ ____ ____ ____ Stimme nicht zu

Ich würde die Sprachsteuerung weiterempfehlen

Stimme zu ____ ____ ____ ____ Stimme nicht zu

Ich möchte gerne eine Sprachsteuerung im Fahrzeug haben

Stimme zu ____ ____ ____ ____ Stimme nicht zu

Freie Beurteilung

Möchten Sie uns noch etwas mitteilen? Haben Sie weitere Anregungen oder Kritikpunkte? Haben Sie Funktionen vermisst? Was fanden Sie besonders gut oder schlecht?

Herzlichen Dank für die Beantwortung des Fragebogens
und für Ihre Teilnahme an unserem Test !!!